VOLUME 107 PART B NUMBER 33 MAY 1960

The Proceedings

OF

THE INSTITUTION OF ELECTRICAL ENGINEERS

FOUNDED 1871: INCORPORATED BY ROYAL CHARTER 1921

PART B

ELECTRONIC AND COMMUNICATION ENGINEERING
(INCLUDING RADIO ENGINEERING)

SAVOY PLACE · LONDON W.C.2

Price Ten Shillings and Sixpence

THE INSTITUTION OF ELECTRICAL ENGINEERS

FOUNDED 1871 INCORPORATED BY ROYAL CHARTER 1921

PATRON: HER MAJESTY THE QUEEN

COUNCIL 1959–1960

President

SIR WILLIS JACKSON, D.Sc., F.R.S.

Past-Presidents

W. H. ECCLES, D.Sc., F.R.S.
THE RT. HON. THE EARL OF MOUNT EDGCUMBE, T.D.
J. M. DONALDSON, M.C.
PROF. E. W. MARCHANT, D.Sc.
H. T. YOUNG.
SIR GEORGE LEE, O.B.E., M.C.
SIR ARTHUR P. M. FLEMING, C.B.E., D.Eng., LL.D.
J. R. BEARD, C.B.E., M.Sc.
SIR NOEL ASHBRIDGE, B.Sc.(Eng.).
SIR HARRY RAILING, D.Eng.
P. DUNSHEATH, C.B.E., M.A., D.Sc.(Eng.), LL.D.
SIR VINCENT Z. DE FERRANTI, M.C.
T. G. N. HALDANE, M.A.
PROF. E. B. MOULLIN, M.A., Sc.D., LL.D.
SIR ARCHIBALD J. GILL, B.Sc.(Eng.).
SIR JOHN HACKING.
COL. B. H. LEESON, C.B.E., T.D.
SIR HAROLD BISHOP, C.B.E., B.Sc.(Eng.), F.C.G.I.
SIR JOSIAH ECCLES, C.B.E., D.Sc.
THE RT. HON. THE LORD NELSON OF STAFFORD.
SIR GORDON RADLEY, K.C.B., C.B.E., Ph.D.(Eng.).
S. E. GOODALL, M.Sc.(Eng.), F.Q.M.C.

Vice-Presidents

O. W. HUMPHREYS, C.B.E., B.Sc. G. S. C. LUCAS, O.B.E., F.C.G.I. SIR HAMISH D. MACLAREN, K.B.E., C.B., D.F.C.*, LL.D., B.Sc. C. T. MELLING, C.B.E., M.Sc.Tech. A. H. MUMFORD, O.B.E., B.Sc.(Eng.).

Honorary Treasurer

E. LEETE.

Ordinary Members of Council

PROF. H. E. M. BARLOW, Ph.D., B.Sc.(Eng.).
C. O. BOYSE, B.Sc.(Eng.).
PROF. M. W. HUMPHREY DAVIES, M.Sc.
SIR JOHN DEAN, B.Sc.
L. DRUCQUER.
J. M. FERGUSON, B.Sc.(Eng.).
D. C. FLACK, B.Sc.(Eng.), Ph.D.
J. S. FORREST, D.Sc., M.A.
R. J. HALSEY, C.M.G., B.Sc.(Eng.), F.C.G.I.
J. B. HIGHAM, Ph.D., B.Sc.
R. A. HORE, M.A., B.Sc.
F. C. MCLEAN, C.B.E., B.Sc.
B. L. METCALF, B.Sc.(Eng.).
J. R. MORTLOCK, Ph.D., B.Sc.(Eng.).
THE HON. H. G. NELSON, M.A.
R. H. PHILLIPS, T.D.
H. V. PUGH.
J. R. RYLANDS, M.Sc., J.P.
G. A. V. SOWTER, Ph.D., B.Sc.(Eng.).
C. E. STRONG, O.B.E., B.A., B.A.I.
D. H. TOMPSETT, B.Sc.(Eng.).

Chairmen and Past-Chairmen of Sections

Electronics and Communications:
M. J. L. PULLING, C.B.E., M.A.
†G. MILLINGTON, M.A., B.Sc.

Measurement and Control:
PROF. A. TUSTIN, M.Sc.
†J. K. WEBB, M.Sc.(Eng.), B.Sc.Tech.

Supply:
J. R. MORTLOCK, Ph.D., B.Sc.(Eng.).
†D. P. SAYERS, B.Sc.

Utilization:
T. E. HOUGHTON, M.Eng.
†R. A. MARRYAT, B.Sc.(Eng.).

Chairmen and Past-Chairmen of Local Centres

East Midland Centre:
D. H. PARRY, B.Sc.
†D. E. LAMBERT, B.Sc.(Eng.).

North-Western Centre:
F. H. HUTCHINSON, M.Eng.
†PROF. F. C. WILLIAMS, O.B.E., D.Sc., D.Phil., F.R.S.

Mersey and North Wales Centre:
T. A. P. COLLEDGE, B.Sc.(Eng.).
†J. COLLINS.

Northern Ireland Centre:
T. S. WYLIE.
†D. S. MCILHAGGER, Ph.D., M.Sc.

Southern Centre:
W. D. MALLINSON, B.Sc.(Eng.).
†G. BISHOP, B.Sc.

North-Eastern Centre:
H. WATSON-JONES, M.Eng.
†A. T. CRAWFORD, B.Sc.

Scottish Centre:
L. A. AKED, M.B.E.
†R. J. RENNIE, B.Sc.

Western Centre:
H. JACKSON, B.Sc.(Eng.).
†R. W. STEEL.

North Midland Centre:
PROF. G. W. CARTER, M.A.
†J. D. NICHOLSON, B.Sc.

South Midland Centre:
G. F. PEIRSON.
†L. L. TOLLEY, B.Sc.(Eng.).

† Past-Chairman

ELECTRONICS AND COMMUNICATIONS SECTION COMMITTEE 1959–1960

Chairman

M. J. L. PULLING, C.B.E., M.A.

Vice-Chairmen

R. J. HALSEY, C.M.G., B.Sc.(Eng.), F.C.G.I. T. B. D. TERRONI, B.Sc. J. A. RATCLIFFE, O.B.E., M.A., F.R.S.

Past-Chairmen

G. MILLINGTON, M.A., B.Sc. J. S. MCPETRIE, Ph.D., D.Sc.

Ordinary Members of Committee

D. A. BARRON, M.Sc.
P. A. T. BEVAN, C.B.E., B.Sc.
J. BROWN, M.A., Ph.D.
PROFESSOR A. L. CULLEN, Ph.D., B.Sc.(Eng.).
L. I. FARREN, M.B.E.
G. G. GOURIET.
COMDR. C. G. MAYER, O.B.E., U.S.N.R.
J. MOIR.
L. J. I. NICKELS, B.Sc.(Eng.).
B. G. PRESSEY, M.Sc.(Eng.), Ph.D.
N. C. ROLFE, B.Sc.(Eng.).
T. R. SCOTT, D.F.C., B.Sc.
C. WILLIAMS, B.Sc.(Eng.).
W. E. WILLSHAW, M.B.E., M.Sc.Tech.
R. C. WINTON, B.Sc.
A. J. YOUNG, B.Sc.(Eng.).

And

The President (*ex officio*)
The Chairman of the Papers Committee.
PROFESSOR H. E. M. BARLOW, Ph.D., B.Sc.(Eng.) (representing the Council).
E. H. COOKE-YARBOROUGH, M.A. (Co-opted Member).
G. H. HICKLING, B.Sc. (representing the North-Eastern Measurement and Electronics Group).
K. F. SANDER, M.A., Ph.D., B.Sc. (representing the Cambridge Electronics and Communications Group).
J. STEWART, M.A., B.Sc. (representing the Scottish Electronics and Measurement Group).
R. FEINBERG, D.Ing., M.Sc. (representing the North-Western Electronics and Communications Group).
R. E. YOUNG, B.Sc.(Eng.) (representing the South Midland Electronics and Measurement Group).

The following nominees:
Royal Navy: CAPTAIN W. D. F. B. MUSPRATT, R.N.
Army: COL. R. G. MILLER, M.A.
Royal Air Force: GROUP CAPTAIN D. W. ROWSON, B.Sc.(Eng.), R.A.F.

MEASUREMENT AND CONTROL SECTION COMMITTEE 1959–1960

Chairman

PROFESSOR A. TUSTIN, M.Sc.

Vice-Chairmen

C. G. GARTON; W. S. ELLIOTT, M.A.

Past-Chairmen

J. K. WEBB, M.Sc.(Eng.), B.Sc.Tech.; H. S. PETCH, B.Sc.(Eng.).

Ordinary Members of Committee

E. W. CONNON, B.Sc.(Eng.), M.Eng.
A. C. LYNCH, M.A., B.Sc.
A. J. MADDOCK, D.Sc.
R. E. MARTIN.
A. NEMET, Dr.Sc.Techn.
S. N. POCOCK.
W. RENWICK, M.A., B.Sc.
G. A. V. SOWTER, Ph.D., B.Sc.(Eng.).
G. F. TAGG, Ph.D., B.Sc.
R. D. TROTTER, B.Sc.(Eng.).
J. H. WESTCOTT, B.Sc.(Eng.), Ph.D.
F. C. WIDDIS, B.Sc.(Eng.).

And

The President (*ex officio*).
The Chairman of the Papers Committee.
G. A. V. SOWTER, Ph.D., B.Sc.(Eng.) (representing the Council).
C. C. BAXENDALE (representing the North-Eastern Measurement and Electronics Group).
E. J. R. HARDY, B.Sc.(Eng.) (representing the North-Western Measurement and Control Group).
D. L. A. BARBER, B.Sc.(Eng.) (nominated by the National Physical Laboratory).
H. M. GALE, B.Sc.(Eng.) (representing the South Midland Electronics and Measurement Group).
W. H. P. LESLIE, B.Sc. (representing the Scottish Electronics and Measurement Group)

Secretary

W. K. BRASHER, C.B.E., M.A., M.I.E.E.

Deputy Secretary

F. JERVIS SMITH, M.I.E.E.

Principal Assistant Secretary

F. C. HARRIS

Editor-in-Chief

G. E. WILLIAMS, B.Sc.(Eng.), M.I.E.E.

TRANSISTORISE *your switching circuitry!*

20 μsec or 40 μsec

2G104 GERMANIUM — **2S101 SILICON**

BVA SO-12

JEDEC TO-18

	P_{tot}		P_{tot}
150 mW	f_1	300 mW	
300 Mc/s		150 Mc/s	f_1
15 V	V_{CB}	25 V	V_{CB}

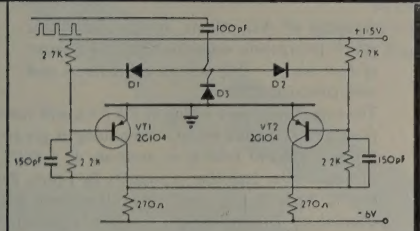
BVA SO-12

JEDEC TO-18

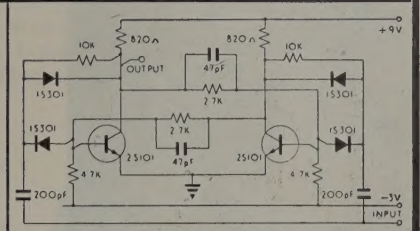
LOW SATURATION (BOTTOMING) VOLTAGE

OPERATION UP TO 150°C

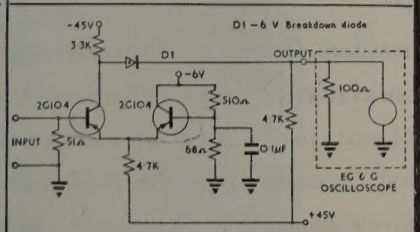
10 Mc/s SATURATED SWITCHING



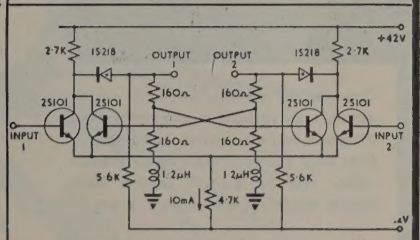
5 Mc/s SATURATED SWITCHING



50 Mc/s NON-SATURATED SWITCHING



25 Mc/s NON-SATURATED SWITCHING



AVAILABLE IN PRODUCTION QUANTITIES NOW

Write for Data Sheets and Application Information on these and other TEXAS devices

I.E.A. EXHIBITION
23rd-28th May.
Olympia
Stand No. E 216

OTHER TEXAS SEMICONDUCTOR DEVICES INCLUDE

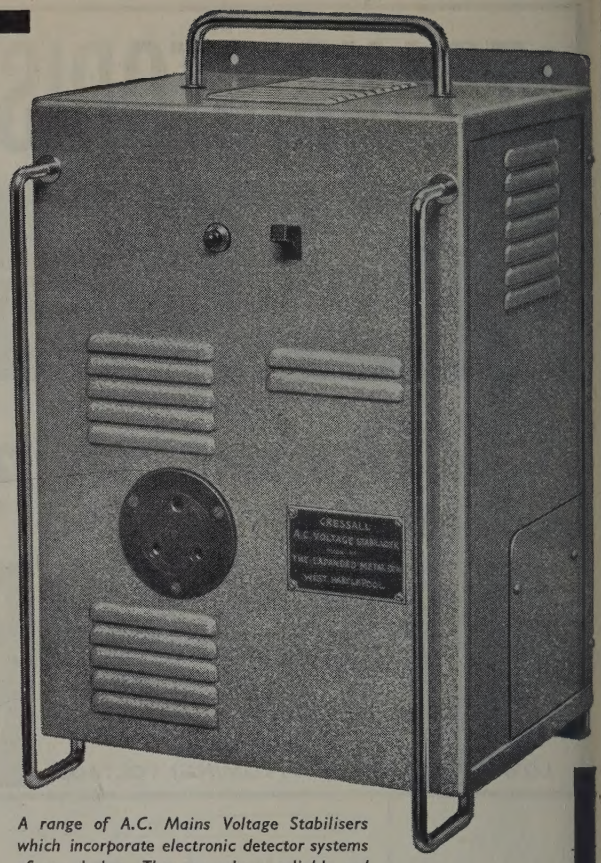
NPN SILICON TRANSISTORS · U.H.F. · Power
NPN GERMANIUM TRANSISTORS · Switching · U.H.F. · Power
SILICON RECTIFIERS & DIODES · High Voltage · Signal · Computer · Photo · Zener
SILICON ZENER VOLTAGE REGULATORS
SILICON CONTROLLED RECTIFIERS
SEMICONDUCTOR-GRADE SILICON

TEXAS INSTRUMENTS LIMITED

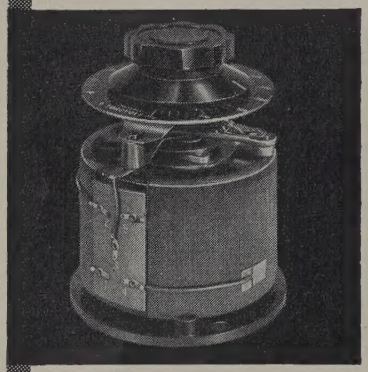
DALLAS ROAD BEDFORD ENGLAND
BEDFORD 68051 CABLES: TEXINLIM · BEDFORD

91/A5

VOLTAGE CONTROL



A range of A.C. Mains Voltage Stabilisers which incorporate electronic detector systems of new design. They are robust, reliable and very compact.
The unit shown has a rating of 2.5 KVA, will handle input variations from 200-250 Volts whilst maintaining a pre-set controlled voltage between 220-240 Volts to an accuracy of ± 0.5%
This particular Stabiliser is suitable for bench, floor or wall mounting and is fitted into a steel case attractively finished in hammered aluminium. Dimensions are : Height 15", Width 10", Depth 7½".



Torovolts are the new toroidally wound continuously variable auto-transformers. Electrical and mechanical improvements have been incorporated to give greater safety and reliability. Legible dials ensure maximum convenience in use.

Expamet-Cressall resistors, rheostats and heaters form the most comprehensive range available from a single source. Units to control loads from 4 watts to over 20,000 Kilowatts are already in use all over the world. To widen still further the enormous field of applications of Expamet-Cressall products, two new designs have been introduced and are briefly described here. Further details and technical advice from our engineers is freely available.

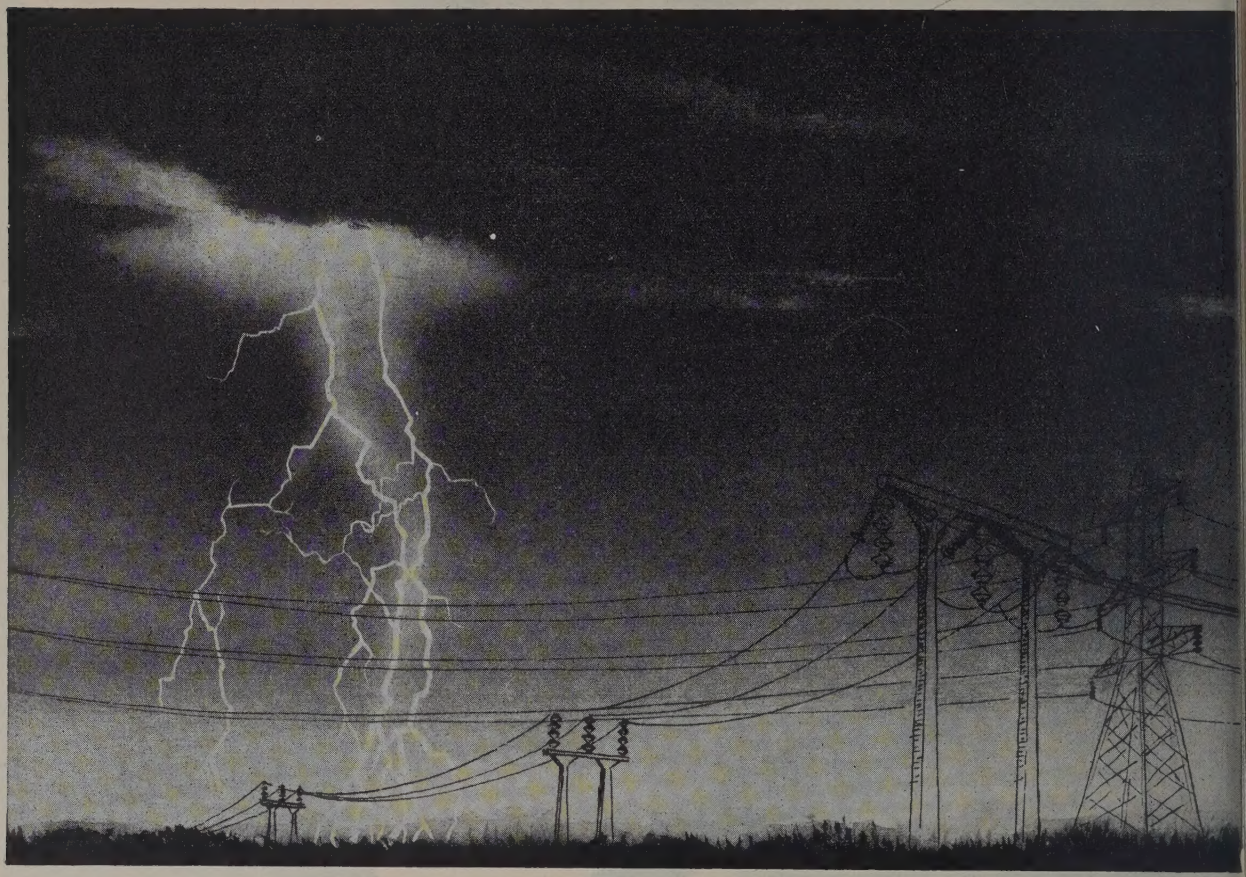
EXPAMET—CRESSALL

Expamet Cressall

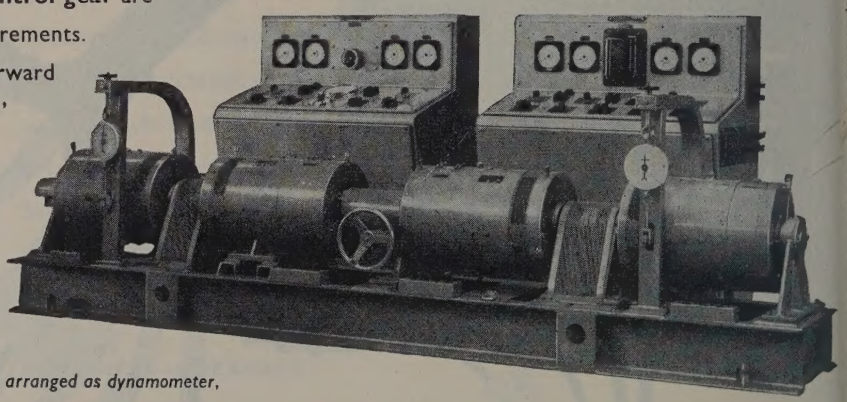
The Electrical Division of
The Expanded Metal Company Ltd.

London Office: Burwood House, Caxton Street, London, S.W.1. Telephone: ABBey 7766
Works: Stranton Works, West Hartlepool. Tel: Hartlepools 5531
The Cressall Manufacturing Company Ltd., Eclipse Works, Tower Street, Birmingham 19
Tel: Aston Cross 2666

POWER AND CONTROL



E.D.C.C. machines and control gear are available to meet most requirements. Standard units for straightforward applications and 'tailor made' specials for the most diverse uses.



Recently supplied to the Welsh College of Advanced Technology, Cardiff—a multiple motor generator set, comprising 4 machines and two control desks. The machine unit comprises:—

- *7 h.p. D.C. variable speed motor, arranged as dynamometer, complete with Tacho-generator*
- *4kVA alternator*
- *4kVA alternator with rotable stator*
- *7 h.p. synchronous motor (left) arranged as dynamometer*

- *The complete unit can be split into two motor-alternator sets*

ELECTRO-DYNAMIC

CONSTRUCTION COMPANY LIMITED

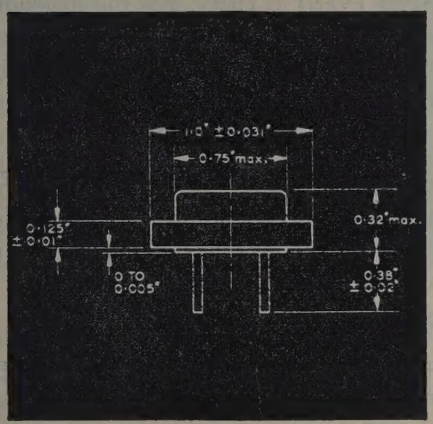
ST. MARY CRAY • ORPINGTON • KENT • TELEPHONE: ORPINGTON 27551 • GRAMS: ELEDAMIC ST. MARY CRAY

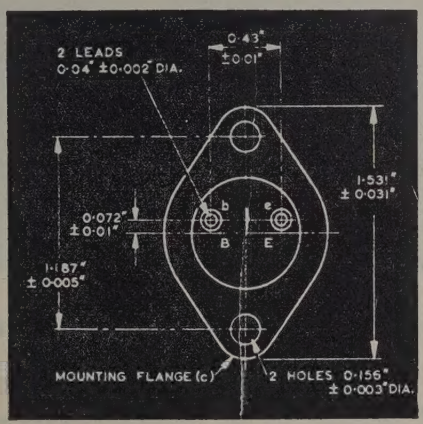
CONTROL GEAR DIVISION: BRIDGWATER SOMERSET TEL: BRIDGWATER 2882 GLASGOW OFFICE: 40 HOULDSWORTH ST C3 TEL: CENTRAL 2620

Audio power output transistors

TYPES XC141 and XC142

These germanium p-n-p alloy junction transistors are designed for use in Class A and Class B power output stages of audio frequency amplifiers. Full particulars of these and other Ediswan Mazda semiconductor devices will be sent gladly on request. If you wish to be kept up to date with the latest developments in this field, please ask us to add your name to our semiconductor mailing list.



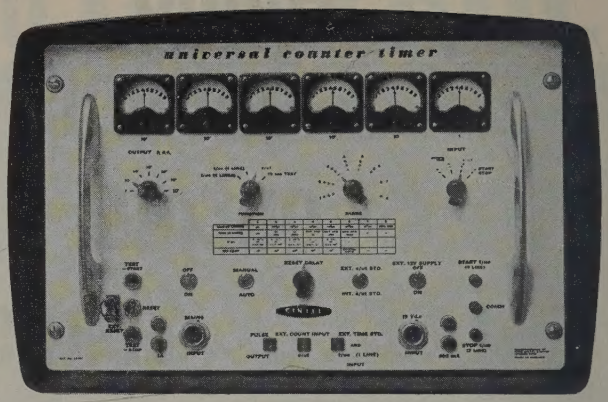


MAXIMUM RATINGS (Absolute Values)	XC141	XC142
Peak collector to base voltage (volts)	−40	−60
Peak collector to emitter voltage, emitter non-conducting (volts)	−40	−60
Peak collector to emitter voltage, emitter conducting (volts)	−32	−32
D.C. Emitter to base voltage (volts)	−12	−12
Peak collector current (amps)	−3·0	−3·0
D.C. Collector current (amps)	−1·5	−1·5
Collector dissipation (mounting flange temperature 80°C) (watts)	11	11

EDISWAN SEMICONDUCTORS
MAZDA

Associated Electrical Industries Ltd
Radio and Electronic Components Division
PD 15, 155 Charing Cross Road, London, W.C.2
Tel: GERrard 8660 Telegrams: Sieswan Westcent London

Transistorized UNIVERSAL COUNTER TIMER



Frequency Measurement

Random Counting

Frequency Division

Time Measurement

Frequency Standard

This fully transistorized portable equipment provides for a wide range of time and frequency measurement as well as facilities for counting, frequency division and the provision of standard frequencies. The facilities available are briefly listed below:

TIME/UNIT EVENT (1 LINE): For the measurement of the time interval between two occurrences in a continuously varying electrical function in the range 3μsec to 1sec. The time for 1, 10 or 100 such events can be measured.

TIME/UNIT EVENT (2 LINE): For time measurement in range 1μsec to 2777hrs. of any interval defined by a positive or negative going pulse in any combination.

EVENTS/UNIT TIME: For frequency measurement in range 30c/s to 1Mc/s over period of 0·001, 0·01, 0·1, 1 or 10secs. Crystal accuracy ±2 parts in 10^6/week. For mains or 12Vd.c. operation.

Full technical specification available on request.

CINTEL INSTRUMENT Division

RANK CINTEL LIMITED
WORSLEY BRIDGE ROAD · LONDON · SE26
HITHER GREEN 4600

Sales and Servicing Agents: Atkins, Robertson & Whiteford Ltd. Industrial Estate, Thornliebank, Glasgow; McKellen Automation Ltd., 122 Seymour Grove, Old Trafford, Manchester, 16; Hawnt & Co. Ltd., 59 Moor St., Birmingham, 4.

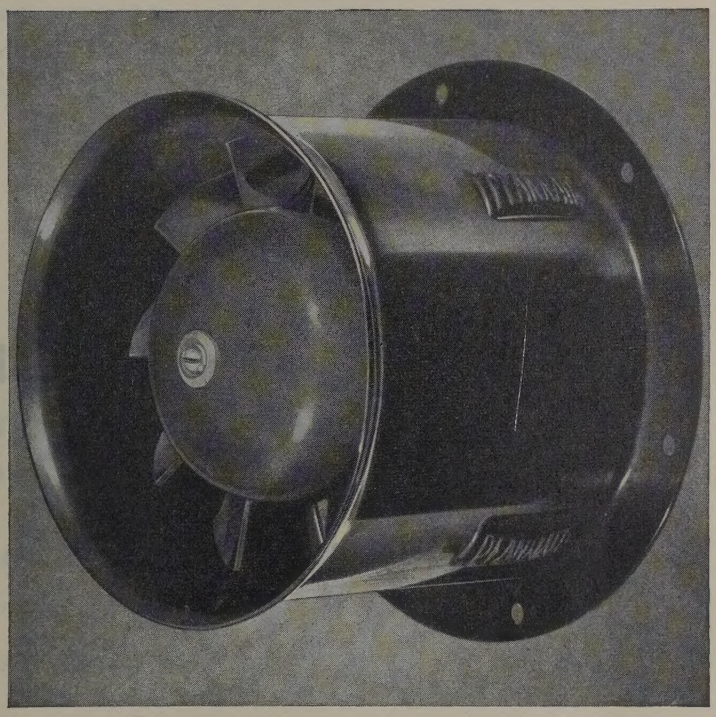
if you had an all seeing eye

If you had an all seeing eye you would at once realise the importance of Plannair Wafters, for you would be able to observe them simultaneously at work in very many electronic equipments and those of allied fields.

At once you would be impressed by their reliability and adaptability to miniaturised electronic equipment, wherever a small general purpose blower is needed.

Plannair Wafters offer the Design Engineer a compact, lightweight, streamlined and particularly silent answer to problems arising from localised 'hot spots'.

Plannair, as a specialising body, is completely single-minded in its dedication to the solving of air movement and temperature control problems. Plannair contributes the most efficient blowers, weight to output, in the world today. Our knowledge could be of assistance, you think? Then we'd be more than happy to help!



The Plannair Wafter Unit 3PL101-409

TECHNICAL SPECIFICATION:

HOUSING Designed to comply with the Ministry of Aviation requirements to meet the Specification No. 1086B for tropicalisation. In light alloy anodised finish incorporating black moulded rubber resilient mouldings for supporting motor. Terminal block mounted on outside.

DIMENSION $3\frac{1}{4}''$ long, 4" dia. over bell mouth, $4\frac{3}{4}''$ dia. over rear flange.

WEIGHT 1 lb. 6 oz.

IMPELLER Moulded fibre glass.

MOTOR Shaded pole type wound for 110/115 V or 230 V single phase 50 cycles. Consumption 12 watts.

PERFORMANCE 50 c.f.m. under free air conditions.

SPEED 2,600 r.p.m.

NOTE: *The Plannair Wafter is available in a number of alternative housings. Send for the Plannair Wafter leaflet which gives full details of units available.*

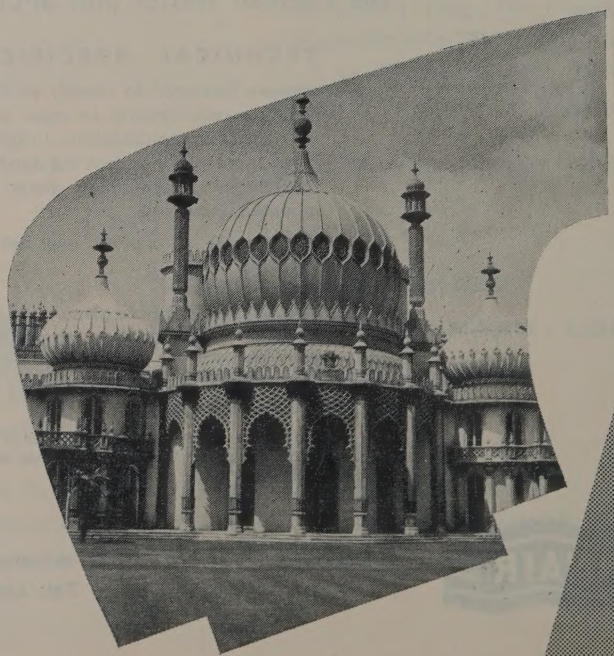
PLANNAIR

PLANNAIR LIMITED Windfield House, Leatherhead, Surrey.
Tel: Leatherhead 4091/3 & 2231

PLA33a

A complete 600 channel four tube repeater

BRIGHTON –CHICHESTER LINK BY A.T.E



est type of small core coaxial system will be installed to
a new 35-mile telecommunication link between Brighton and
ter. It will provide the G.P.O. with an economic, high
system for circuits between the two towns.
ried type transistorised repeater is of a completely new design
ed by A.T.E. It is power-fed along the cable and contained in
etically sealed, pressurised nickel iron box, suitable for
tion in footway boxes or underground chambers. For full
write for publication T.E.B 3202.

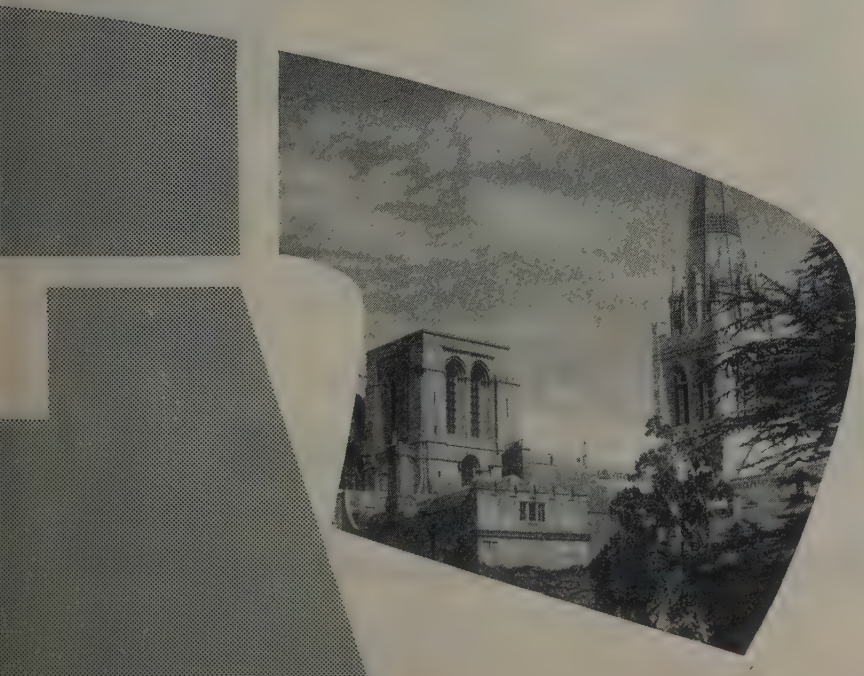
ly transistorised system of proved reliability
ts latest C.C.I.T.T. recommendations
to 300 channels per pair of tubes
npact power-fed repeaters
built supervisory and maintenance facilities

T.E. Coaxial system type C300A

for use on small core coaxial cables

AUTOMATIC TELEPHONE
& ELECTRIC CO. LTD.,

Strowger House, Arundel Street, London, W.C.2
Tele: TEMple Bar 9262.



AT 8911

Telegraph distortion measurement

in the speed range 40-400 bauds

Rapidly increasing international use of high-speed synchronous telegraph systems has led to the urgent need for suitable distortion measuring equipment. M.S. Type TCA1219, manufactured by A.E.I. at their Woolwich factory to a British Post Office design, enables distortion measurements to be made within 1% accuracy. The method of indication is by bright spots on the circular time base of a cathode-ray tube. Screening and the fitting of radio interference suppressors have been given particular attention.

Size	17 in. high x 15 in. wide x 18 in. deep
Weight	110 lb.
Power supply	200/250 v. 50 c.p.s.
Power consumption	170 w.

vast experience **in the science of telecommunications**

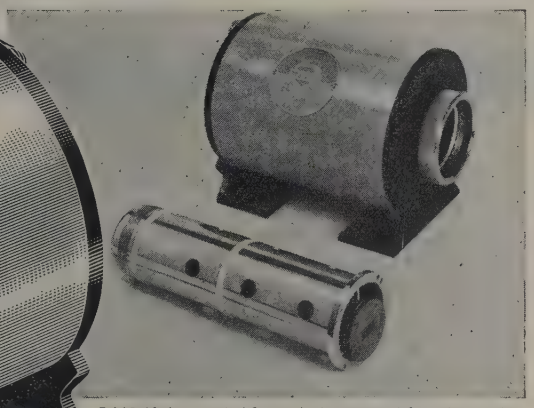
AEI **Telecommunications Division**
Transmission Department Woolwich London SE18
Associated Electrical Industries Limited

TA 3775

Backward Wave Oscillators

Mullard now have available a range of O-type backward wave oscillators covering operation in the S, X and J bands.

The mechanical design allows substantial cuts in maintenance costs to be effected as the focusing system may be re-employed should the valves need to be changed. The valves are supplied pre-aligned in a protective capsule which automatically locates in the focusing system and any replacements may be made quickly and easily without need for focusing adjustments. Both electro-magnet and permanent magnet focusing systems are available.



BA16-10 demounted from electro-magnet focusing system

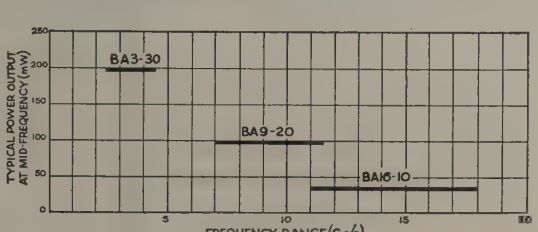
See Mullard valves and tubes on Stand M559 at the INSTRUMENTS, ELECTRONICS AND AUTOMATION EXHIBITION

Where modulation of the valve output is required this can be readily achieved by modulating the appropriate electrode. The output connection is isolated from the delay structure so that the valve may be operated with its cathode earthed and consequently high modulation frequencies may be used.

These specialised microwave valves make possible the design of wide frequency range microwave instruments, microwave search receivers and f.m. carrier systems. Write to the address below for full details of these and other Mullard microwave valves.

Abridged data

Type No.	Frequency Range (Gc/s)	Power Output over Frequency Range (mW)	Delay Structure Voltage Range (V)	Sensitivity at Mid-frequency (Mc/s per V)	Cathode Current max. (mA)
BA3–30	2.4 to 4.5	30 to 500	150 to 1500	3.0	50
BA9–20	7.0 to 11.5	20 to 180	250 to 1400	5.0	28
BA16–10	11 to 18	10 to 70	500 to 2500	3.5	13



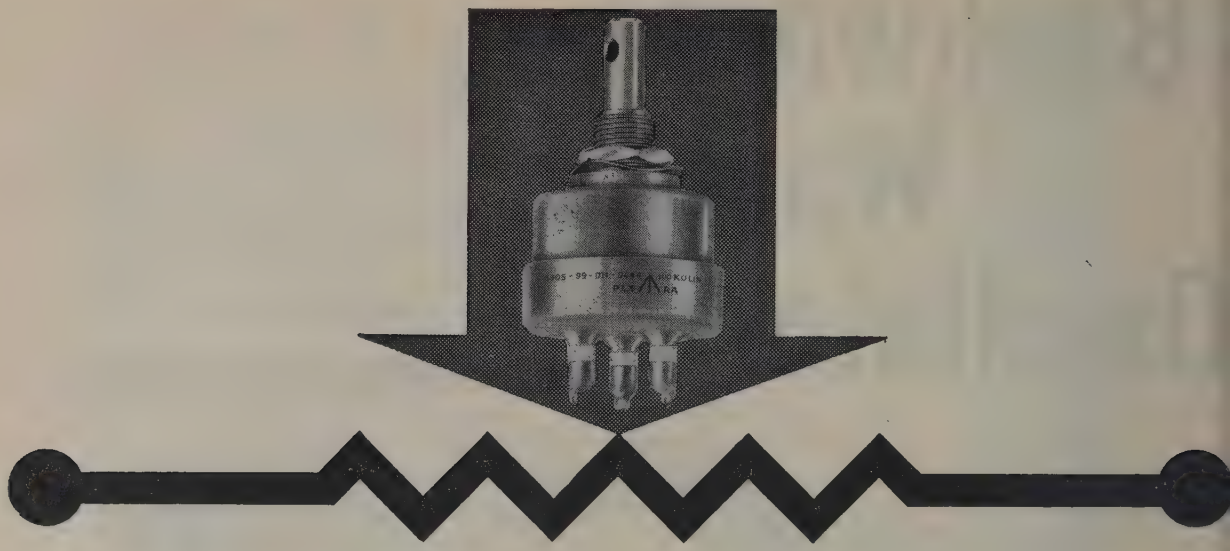
Mullard

Mullard

GOVERNMENT AND INDUSTRIAL VALVE DIVISION

MULLARD LIMITED
Mullard House, Torrington Place, London, W.C.1
Telephone: LANgham 6633

MVT 383



PROOF
against all conditions

The Plessey XP5 *hermetically sealed* moulded carbon track potentiometer

POWER RATING
—¾ watt continuous at 70°C

RESISTANCE RANGE
—500 Ω to 2.5 M Ω

RESISTANCE LAW
—Linear or Logarithmic

INSULATION
—Not less than 5000 M Ω at 500 V.d.c. (applied for 1 minute between spindle and all terminations)

Under the most arduous climatic or service conditions the XP5 will give trouble-free service. Completely sealed against moisture, it will withstand the most severe conditions of bumping, vibration, humidity and tropical exposure. Operating over a temperature range of −40°C to 70°C it will give smooth, noise-free and reliable service—with negligible wear. It is approved to Inter Service Standard RCS 122A.

For full details of this outstanding new component send for leaflet 312.

Overseas Sales Organisation
PLESSEY INTERNATIONAL LIMITED
Overseas Telegrams: PLESSINTER · TELEX · ILFORD
Head Office: Ilford · Essex · England · *Telephone:* Ilford 3040 · *Telex:* 23166 Plessey Ilford
Telegrams: PLESSEY TELEX ILFORD

THE PLESSEY COMPANY LIMITED
Capacitors & Resistors Division · Swindon Group
KEMBREY STREET · SWINDON · WILTSHIRE
Telephone: Swindon 6211 · *Telex:* 44-355 · *Telegrams:* Plessey Telex Swindon
CR6A

STC

proved reliability

CAPACITORS

The use of STC Capacitors in vital communication and navigation equipment throughout the world acknowledges their high degree of reliability. Equipment which must operate with unfailing efficiency over very long periods demands critical performances from every component. This standard of performance is met by STC Capacitors. The production techniques for these special applications were evolved from established processes for all STC Capacitors—a guarantee of consistency and efficiency in *your* equipment.

- ***Extensively life tested to establish long-life performance.***
- ***Manufactured in closely controlled atmospheres.***
- ***Process controlled at all stages of production.***
- ***Tested on equipment specially designed to simulate operating conditions.***

Illustrated is a repeater station in the new Anglo-French microwave telephone and television link. STC Capacitors are used in this and many other SHF communications systems throughout the world.

Write for STC Capacitor literature to:

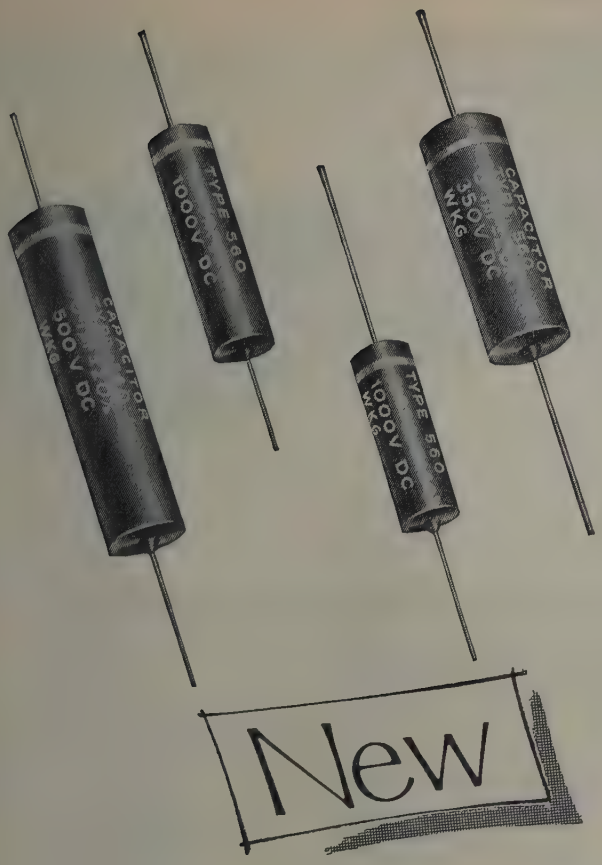
STC COMPONENTS GROUP

Standard Telephones and Cables Limited

Registered Office: Connaught House, Aldwych, London, W.C.2

CAPACITOR DIVISION: BRIXHAM ROAD · PAIGNTON · DEVON

60/1MC



- No exposed metal parts other than terminations, which are clean solder coated, thereby ensuring easy soldering.
- Body and terminations free of wax coating or any other low melting point material.
- Long life without voltage derating.
- Designed to meet the requirements of British Joint Service Standards RCS 131 and BS 2131 with humidity classification H.2.
- Solid construction eliminates internal movement, preventing damage by severe vibration.

DUBILIER ENCAPSULATED PAPER DIELECTRIC TUBULAR CAPACITORS HAVING OUTSTANDING CHARACTERISTICS

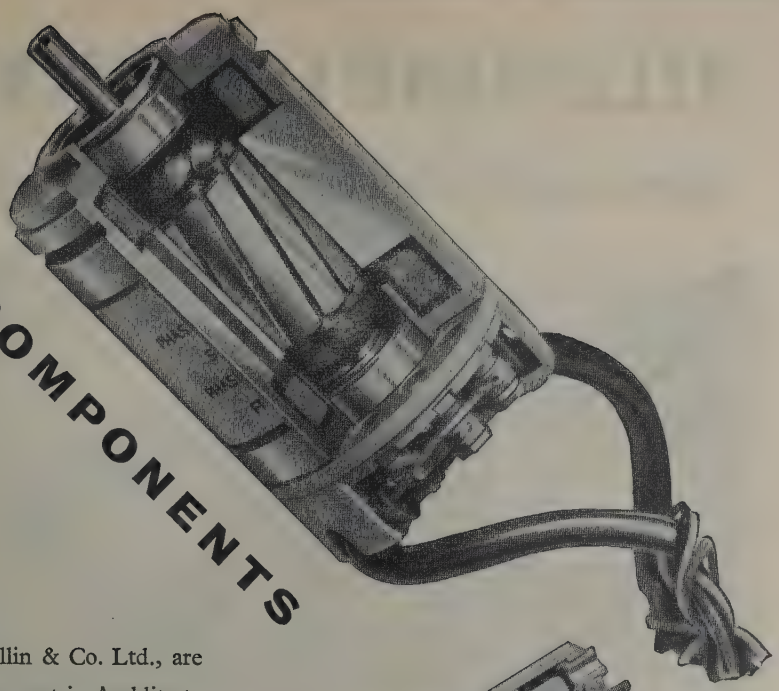
The Dubilier Capacitor Type 560 is a new approach to capacitor requirements for all radio and electronic applications. It is constructed to meet long and arduous service conditions. The paper dielectric element is impregnated with a plastics material to produce a solid unit. The terminations are of great mechanical and electrical strength and the assembled element is sealed in an encapsulated mineral loaded epoxy resin so that there are no parts capable of movement, making the capacitor completely immune to shock and all normal atmospheric conditions.
Capacitance Tolerance; ±20% normal ±10% by selection. Power Factor; Less than 1% at 1,500 c/s. Insulation Resistance; Better than 20,000MΩ at normal temperature. Voltage Application; From—40° to +125°C for d.c. and from —40° to +70°C for a.c.

CAPACITANCE µF	VOLTAGE RATINGS			DIMENSIONS	
	d.c. Wkg. at —40°C to +125°C	d.c. Test at 20°C	a.c. Wkg. r.m.s. at —40°C to +70°C and up to 60 c/s	Diameter +0.020″ −0	Length ±0.040″
0.001	1,000	2,500	250	3/8	1
0.002	1,000	2,500	250	3/8	1
0.005	1,000	2,500	250	3/8	1
0.01	1,000	2,500	250	3/8	1 3/8
0.02	750	2,250	250	3/8	1 3/8
0.05	500	1,500	250	1/2	1 3/8
0.1	350	1,000	180	1/2	1 3/8
0.1	500	1,500	250	1/2	1 15/16

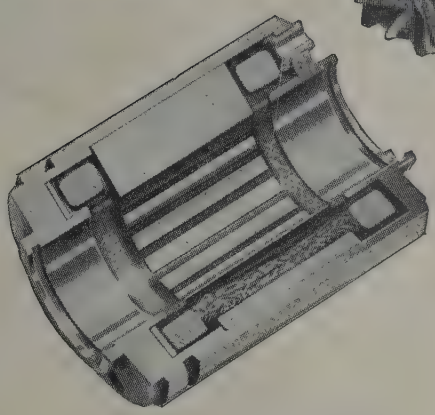
DUBILIER

DUBILIER CONDENSER CO. (1925) LIMITED, DUCON WORKS, VICTORIA ROAD, NORTH ACTON, LONDON W.3
Telephone: ACOrn 2241 (5 lines)
Cables: Hivoltcon London
Telegrams: Hivoltcon London Telex
Telex: 25373
DN 242A

Araldite IN SERVOCOMPONENTS

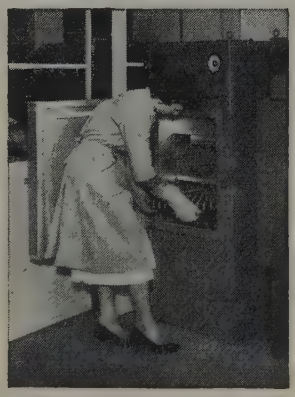


The Synchro units shown, made by R. B. Pullin & Co. Ltd., are sectioned to show how the stators are integrally cast in Araldite to provide maximum protection against the effects of extremes of temperature, humidity and vibration. The excellent machining properties of Araldite make possible a straight-through bore technique, which eliminates errors in alignment and also permits the use of the smallest possible air gap between rotor and stator. High insulation and dielectric strength, remarkable adhesion to metals, and negligible shrinkage on curing make Araldite eminently suitable for use in the construction of precision electrical equipment.



Araldite epoxy resins are used—

- **for casting high grade solid electrical insulation**
- **for impregnating, potting or sealing electrical windings and components**
- **for producing glass fibre laminates**
- **for producing patterns, models, jigs and tools**
- **as fillers for sheet metal work**
- **as protective coatings for metal, wood and ceramic surfaces**
- **for bonding metals, ceramics, etc.**



This photograph shows an A.E.W. electric oven, capable of maintaining temperatures within close limits, as used by R. B. Pullin & Co. Ltd. for curing the Araldite-filled stators.

Araldite *epoxy resins*

Araldite is a registered trade name

CIBA (A.R.L.) LIMITED

Duxford, Cambridge Telephone: Sawston 2121

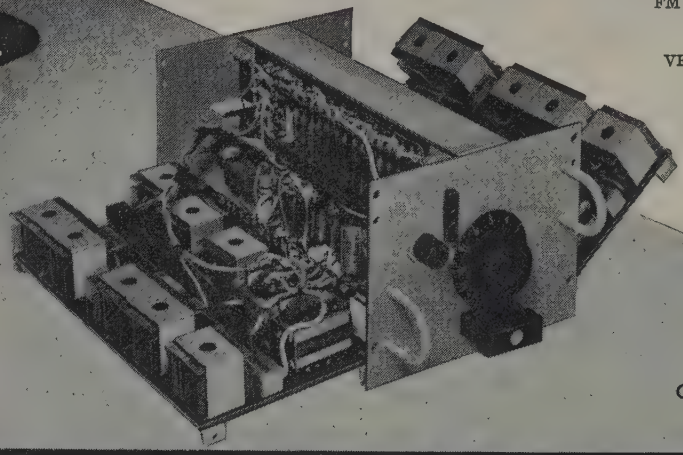
AP 463

TELEPHONE EQUIPMENT BY T.M.C.

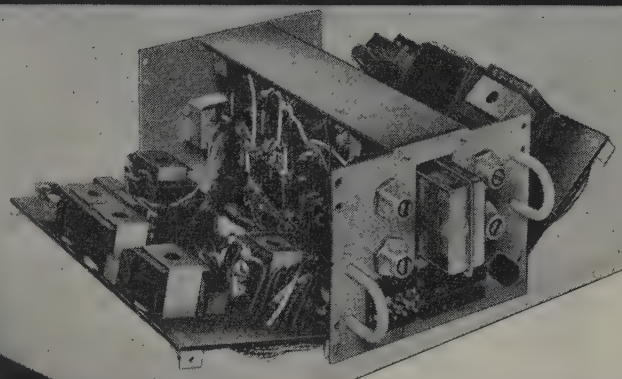
Channelling Equipment

This is the latest addition to the T.M.C. range which includes: 2, 3, 4 and 6 kc/s spaced carrier telephone equipment for cable, radio and open-wire systems. 120, 170 and 240 c/s spaced FM telegraph equipment. Compandors, negative impedance repeaters, VF amplifiers and privacy equipment.

a



Channel Unit



Signalling Unit

b

GENERAL

- As multi-channel carrier telephone equipment, it is designed to provide high-quality wide-band telephone circuits.
- It can be energized from suitable batteries or a.c. mains.
- Plug-in units and panels make transport, installation and maintenance easier.
- Great care has been taken to simplify the problem of installation. With plug-in panels and units removed, the rackside is easily lifted into position. Racksides may be mounted side-by-side back-to-back or back to a wall. Cable entries are arranged for both overhead and floor-duct distribution.

This new equipment is a further T.M.C. contribution to the recent advances made by the electronics industry in the field of telecommunication. In comparison with similar equipment, it effects considerable economies in the reduction of space used and power consumed. And still more economies are brought about by its simplicity of installation and ease of maintenance.

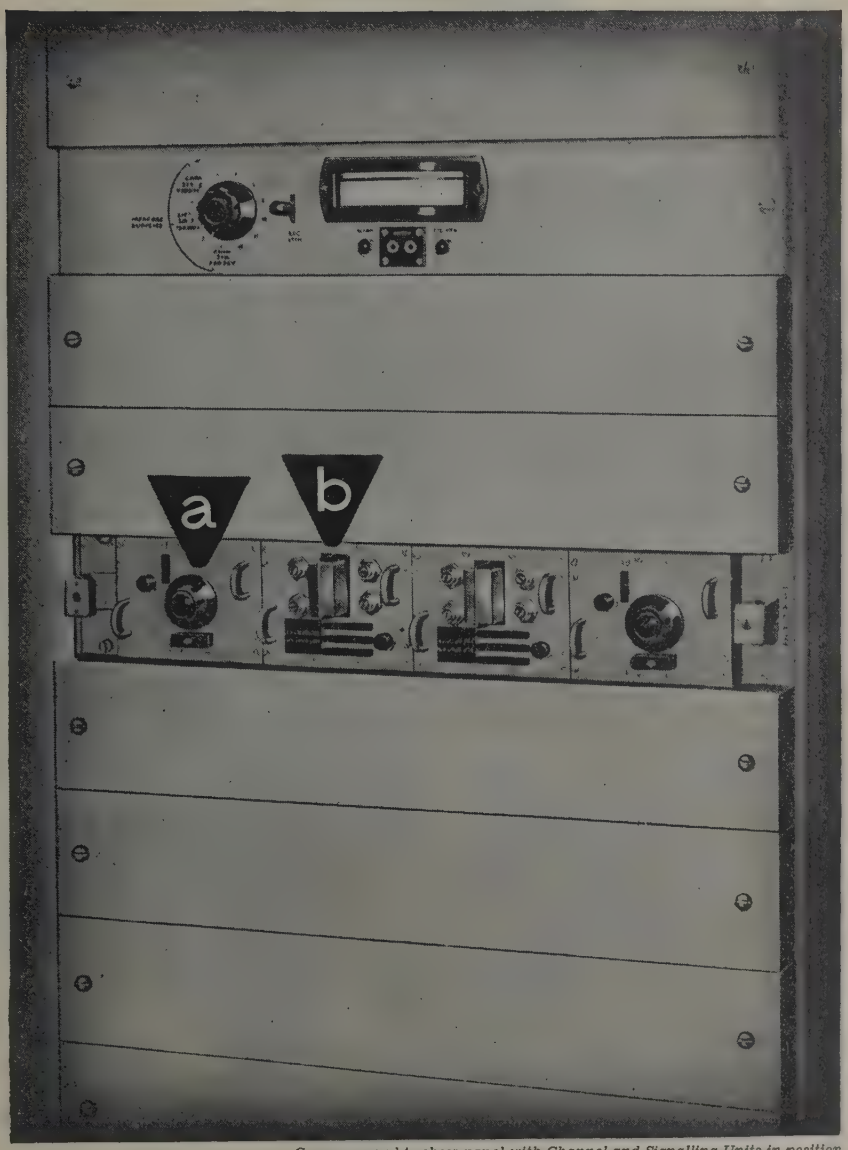
TMC

NEW TRANSISTORIZED CARRIER

For Cable Systems · Radio Systems ·

TECHNICAL FEATURES

- 300-3,400 c/s Channel Bandwidth.
- 4 kc/s spaced.
- Optional Inbuilt Out-band signalling (3825 c/s at -20 dbmO)
- Fully transistorized.
- 96 Circuits without signalling or 48 circuits with signalling per 9 ft. rackside (20½ in. x 8½ in. floor dimensions).
- Compact plug-in units with hinged, card-mounted components give maximum component accessibility.
- Can be supplied on racksides or a complete basic group on a sub-frame.
- Available in 60-108 kc/s basic channel groups or as systems for 12-circuit 2-wire cable operation or multi-circuit, for radio links.
- Conforms with CCITT and BPO requirements.



Cover removed to show panel with Channel and Signalling Units in position

TMC

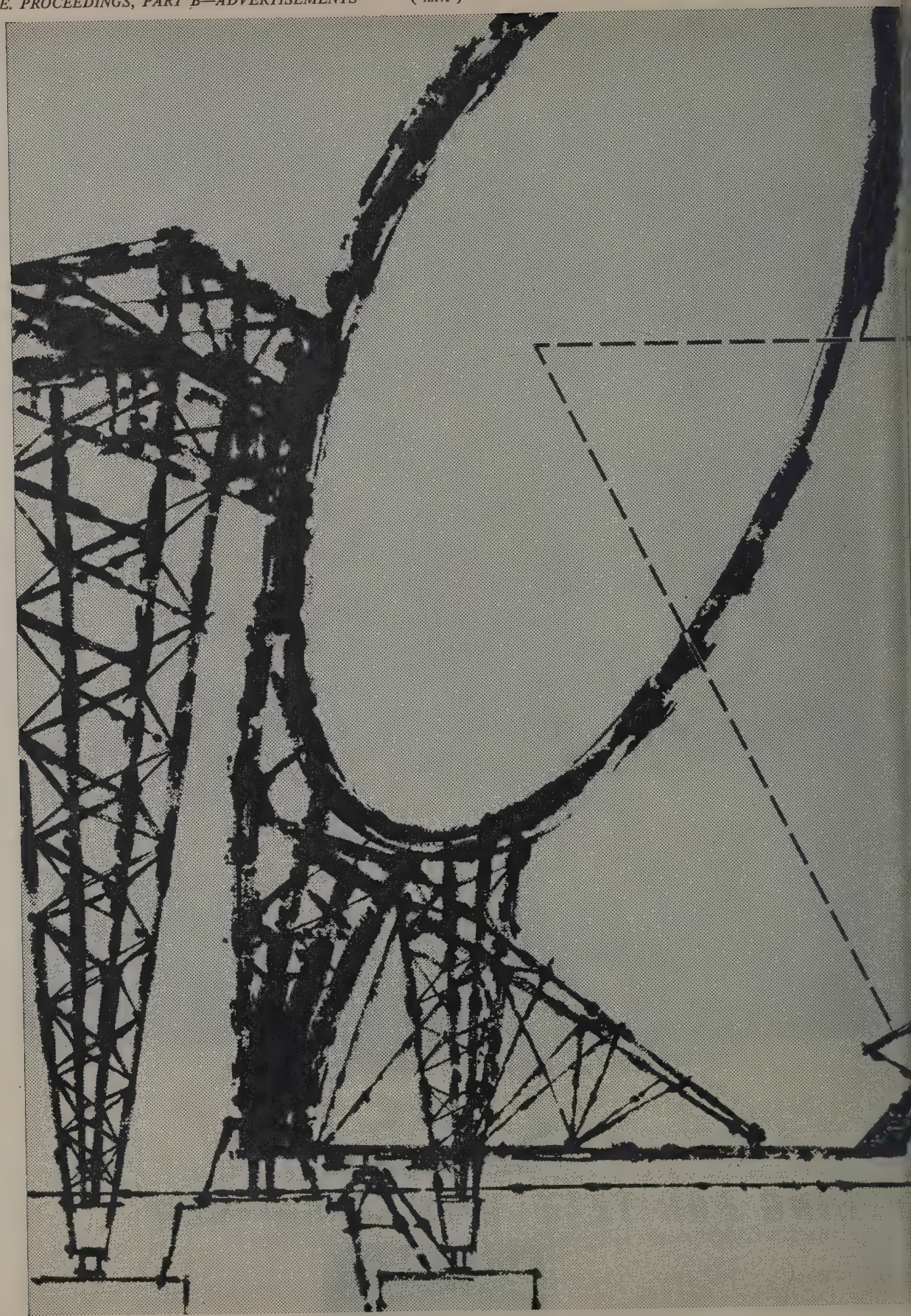
Further information from:

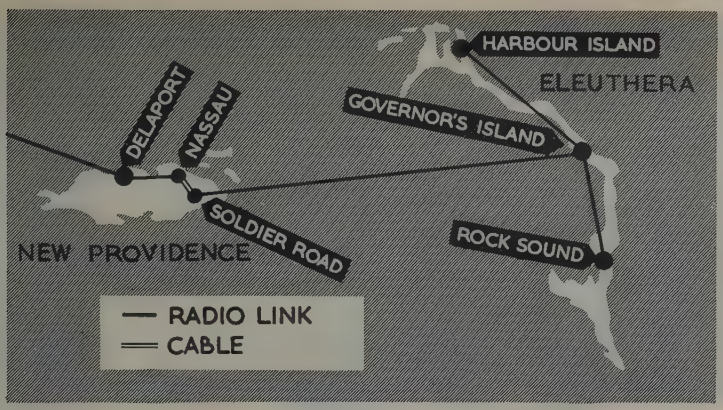
TELEPHONE MANUFACTURING COMPANY LIMITED

Transmission Division: Cray Works, Sevenoaks Way, Orpington, Kent

Telephone: Orpington 26611

SELLING AGENTS *Australia and New Zealand:* Telephone Manufacturing Co. (A'sia) Pty Ltd., Sydney, New South Wales.
Canada and U.S.A.: Telephone Manufacturing Co. Ltd., Toronto, Ontario.
All other Countries (for transmission equipment only): Automatic Telephone and Electric Co. Ltd., London.





STC O.T.H. in the Bahamas

STC High Power, forward-scatter radio telephone equipment installed near Nassau is now operating as one end of a 72 circuit link which connects the Bahamas with the U.S.A. nation-wide telephone network. This STC equipment is working in conjunction with U.S. equipment installed in Florida. Signals are transmitted and received over a single 180 mile over-water path.

Coupled with Microwave and V.H.F. Radio Links

The O.T.H. installation at Delaport is connected to the telephone exchange in Nassau City by a 7400 Mc/s line-of-sight microwave link. The flourishing tourist centres in the Eleuthera Islands are connected to Nassau City by a V.H.F. radio-telephone network comprising 3 links.

STC have also supplied:

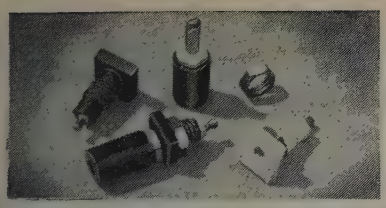
a 100 kVA diesel alternator standby power machine for the O.T.H. terminal;
frequency multiplexing equipment;
4 telephone exchanges with a total of 22 switchboard positions.

ELECTRONIC SYSTEMS GROUP

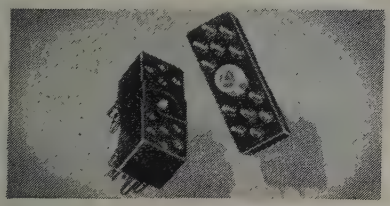
Standard Telephones and Cables Limited

Registered Office: Connaught House, Aldwych, London, W.C.2

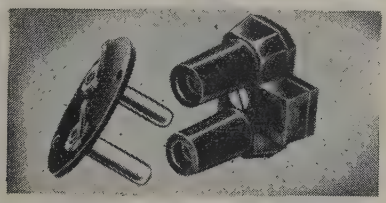
TRANSMISSION DIVISION: NORTH WOOLWICH · LONDON · E.16

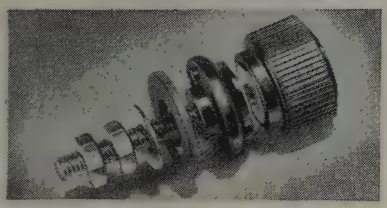


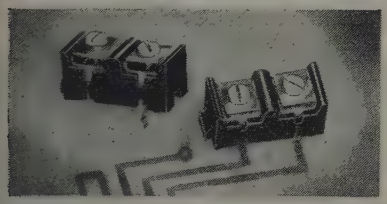
"BELLING-LEE" CONNECTORS

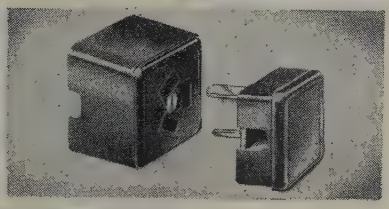


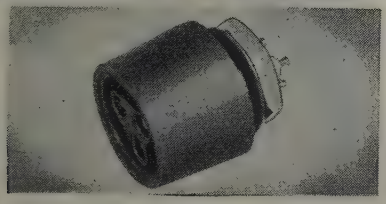
Bricks for the Electronic Industry

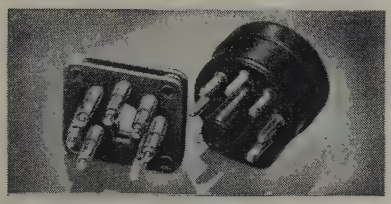












The "BELLING-LEE" range of connectors is probably the most extensive in the industry, catering for a most comprehensive variety of applications.

The latest addition is a Miniature Mains Plug-and-socket, with a rating of 1 amp., and occupying no more panel space than a half-penny. Moulded in nylon with fully shrouded plug pins, this space-saving accessory will stand a surprising amount of rough treatment. Although miniature in size, it is a giant in performance.

Please write for details of this useful component, and if you have any problem concerning connections why not let us help you to solve it? Over 30 years' intensive experience of design and manufacturing in this field is yours for the asking.

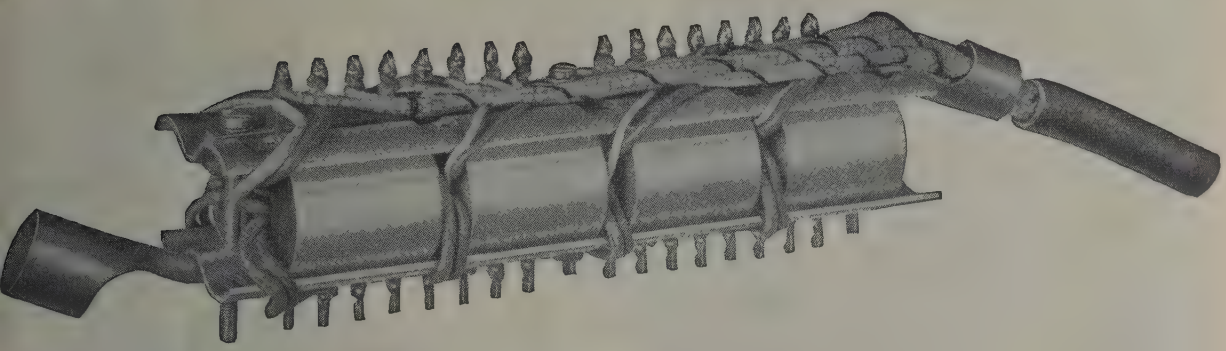
STAND D165 · I.E.A. EXHIBITION

TERMINALS · PLUGS & SOCKETS · GLASS SEALS
CIRCUIT PROTECTION DEVICES
INTERFERENCE FILTERS · RECEIVING AERIALS

Most "Belling-Lee" products are covered by patents or registered designs or applications.

BELLING & LEE LTD
GREAT CAMBRIDGE ROAD, ENFIELD, MIDDX., ENGLAND
Telephone: Enfield 3322 · Telegrams: Radiobel, Enfield

Improved splice loading with the L219

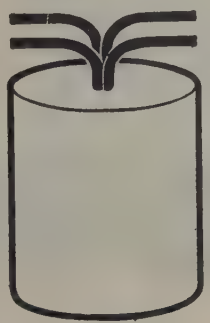


The new economical loading coil

Arising from the increasing demand for a smaller coil which can be employed in splice loading, the L219 has been developed. In the design Mullard Equipment Limited were assisted by their own production experience and information given by overseas users. The result is a simple, low cost component (to grade 3 spec.) suitable for small or large splice loading units.

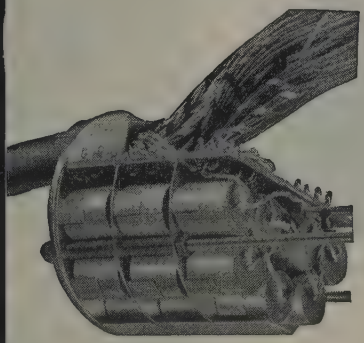
Smaller construction

By using a new grade of Ferroxcube pot core the overall volume of the coil is considerably reduced. The coil is resin sealed in a small cylindrical aluminium canister ensuring complete protection from climatic effects. The windings of the coils are brought out on flying leads.



LIFE SIZE COIL
L219

Permits smaller splices



Key factors in this development are the clamping arrangements which, with the new coil, permit much smaller splice housing. On small cables, coils are mounted lengthways in pairs with great compactness. For larger cables, coils are mounted radially, each mounting plate accommodating up to seven coils. Clamping plates, coils, etc. can be supplied as kits.

Please write for full details of these new loading coils

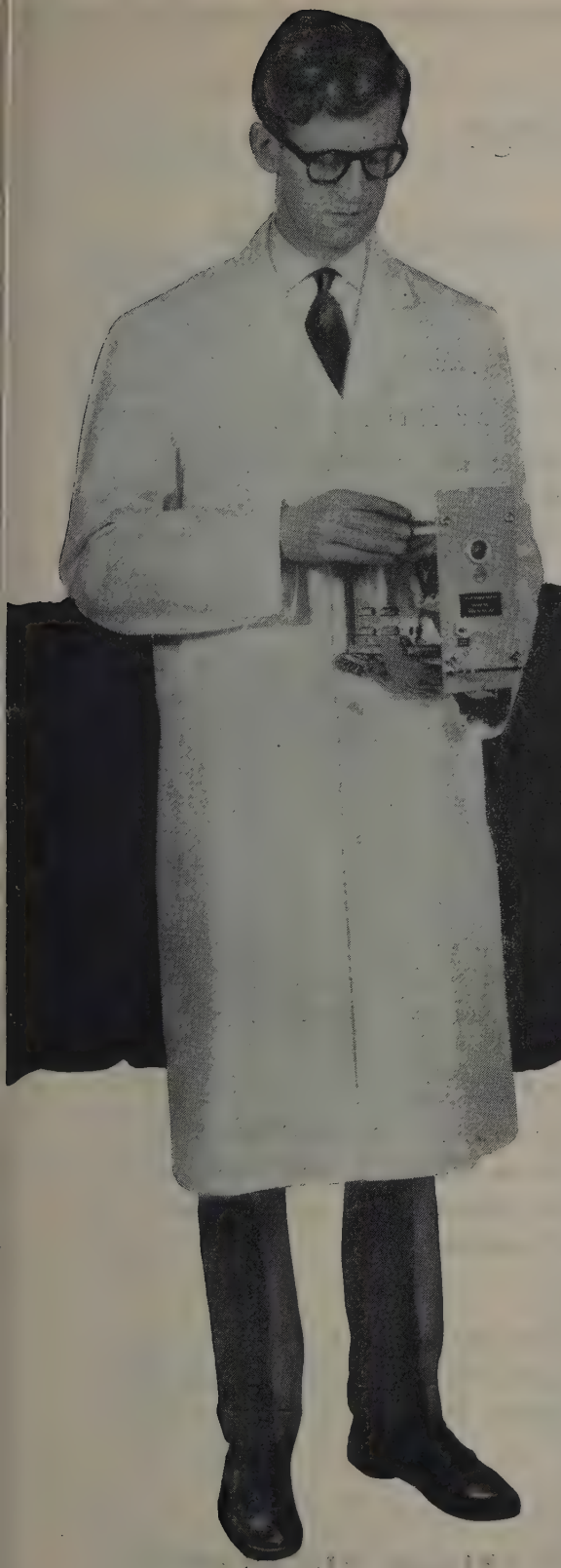
PRODUCTS OF
MULLARD EQUIPMENT LIMITED
A COMPANY OF THE MULLARD GROUP

MULLARD EQUIPMENT LIMITED

A Company of the Mullard Group

Mullard House · Torrington Place · London W.C.1 · Telephone: Langham 6633

ME614a



These new STC complete packaged oscillator units represent a revolutionary step forward in crystal design. They are extremely small units for use as laboratory references or as the oscillator section of high performance equipment. The ruggedness of design together with the small size makes possible a degree of portability hitherto unknown in this type of unit. Standard frequency is 5 Mc/s.

Write for Technical Data Sheets.

1 in 10^9 per day frequency standard for laboratories at £195

One of a new, low priced range of frequency standards from STC.

- **Rugged**
- **Compact**
- **Light-weight**
- **Reliable**
- **Low Cost**

COMPONENTS GROUP

60.1MQ

Standard Telephones and Cables Limited

Registered Office: Connaught House, Aldwych, W.C.2.

QUARTZ CRYSTAL DIVISION: HARLOW · ESSEX

INDEX OF ADVERTISERS

Firm	page
cola Products Ltd.	
mec Ltd.	
W. G. Armstrong Whitworth Aircraft Ltd.	
row Electric Switches Ltd.	
sociated Electrical Industries Ltd.	v, xi, xii, xxxv, xliv
tomatic Telephone & Electric Co. Ltd.	viii + ix
ling and Lee Ltd.	xxix
okhirst Igranic Ltd.	
ble Makers Association	
thodeon Crystals Ltd.	xxx
ba (A.R.L.) Ltd.	xix
whurst and Partner Ltd.	
movan Electrical Co. Ltd.	
bilier Condenser Co. Ltd.	xvii
inburgh University Press	xxxvi
co Electronics Ltd.	xxxviii
ctro Dynamic Construction Co. Ltd.	iv
GA Products Ltd.	
iott Bros. Ltd.	xlii
glish Electric Co. Ltd.	
glish Electric Valve Co. Ltd.	xvi
ie Resistor Co. Ltd.	x
R.G. Industrial Corp. Ltd.	xxxii + xxxvi
icssons Telephones Ltd.	
panded Metal Co. Ltd.	ii
rranti Ltd.	xviii
X. Fox Ltd.	
neral Electric Company Ltd. (M.O. Valves)	
neral Electric Company Ltd. (Semiconductors)	xxxiii
neral Electric Company Ltd. (Telecommunications)	xxiv + xxv
orge Kent Ltd.	
Lodge Plugs Ltd.	
London Electric Wire Co. and Smiths Ltd.	xxxvi
Claude Lyons Ltd.	xxxii
Marconi Instruments Ltd.	xxxvii
Marconi Wireless Telegraph Ltd.	xxiii, xli
Metropolitan Plastics Ltd.	
M-O Valve Co. Ltd.	
Mullard Ltd. (Equipment)	xxxi
Mullard Ltd. (Components)	xxviii
Mullard Ltd. (Valves)	xiii
Multicore Solders Ltd.	xliii
Newmarket Transistors Ltd.	
Newton Bros. (Derby) Ltd.	xxxii
Oliver Pell Control Ltd.	
Philips Electrical Ltd.	
Plannair Ltd.	vii
Plessey Co. Ltd. (Swindon)	iii + xiv
Racal Engineering Ltd.	
Rank Cintel Ltd.	vi
Richard Thomas & Baldwins Ltd.	xxxiv
Salford Electrical Instruments Ltd.	
Savage Transformers Ltd.	
Semiconductors Ltd.	xxii
Servomex Controls Ltd.	
Standard Telephones and Cables Ltd.	xv, xxvi + xxvii, xxxix
Texas Instruments Ltd.	i
Telephone Manufacturing Co. Ltd.	xx + xxi
Ultra Electronics Ltd.	
Whiteley Electrical Radio Co. Ltd.	xxx
Zenith Electric Co. Ltd.	xxx

THE PROCEEDINGS OF
THE INSTITUTION OF ELECTRICAL ENGINEERS

EDITED UNDER THE SUPERINTENDENCE OF W. K. BRASHER, C.B.E., M.A., M.I.E.E., SECRETARY

VOL. 107. PART B. No. 33. MAY 1960

621.317.76 : 621.373.5

The Institution of Electrical Engineers
Paper No. 3002 M
Aug. 1959

FREQUENCY VARIATIONS OF QUARTZ OSCILLATORS AND THE EARTH'S ROTATION IN TERMS OF THE N.P.L. CAESIUM STANDARD

By L. ESSEN, O.B.E., D.Sc., F.R.S., Associate Member, J. V. L. PARRY, M.Sc., and J. McA. STEELE, B.Sc.(Eng.).

(The paper was first received 16th December, 1958, and in revised form 1st April, 1959. It was published in August, 1959, and was read before a joint meeting of the MEASUREMENT AND CONTROL SECTION and the ELECTRONICS AND COMMUNICATIONS SECTION 1st December, 1959, and the MERSEY AND NORTH WALES CENTRE 1st February, 1960.)

SUMMARY

The variations in the frequencies of quartz oscillators compared with that of a caesium standard are described and discussed. It is shown that the frequencies of the ring-type oscillators at the N.P.L. and other laboratories do not depart from those given by a linear law by more than ± 5 parts in 10^9 over a period of three years. The quartz oscillators calibrated by the caesium standard have been used to study the variations in the period of rotation of the earth. In addition to an annual seasonal variation of approximately ± 8 parts in 10^9, there was a fairly steady retardation amounting to 1 part in 10^8 in the period between September, 1955, and January, 1958.

(1) INTRODUCTION

Quartz oscillators are widely used as the working standards of time and frequency, and the practice is likely to continue for many years. A detailed and accurate study of their performance is therefore of interest, and this has been made possible for the first time by the introduction of the caesium atomic standard. Previous studies have been limited by the uncertainty in the value of the unit of time which results from two fundamental obstacles in its determination, namely

(*a*) The errors of astronomical observations are such that the results must be averaged over an interval of about 100 days to give an accuracy of 1 part in 10^9.

(*b*) During such an interval, the frequencies of the quartz oscillators and the period of rotation of the earth both change by a significant amount.

The unit obtained is therefore an average value for the interval over which the observations are extended, and frequencies even to the accuracy of 1 part in 10^9 can be given only as average values. It is, however, known from the intercomparisons of quartz oscillators, which can be made with great precision, that the oscillators in general behave in a very uniform manner, and that at once their laws of variation have been determined, they can be used for interpolation and extrapolation with high accuracy. They enable the unit of time at any instant to be given with an error very little in excess of that of the average value.

The combination of a quartz oscillator and an integrating mechanism which counts the number of vibrations and transforms it into seconds, minutes and hours is known as a quartz clock. The rate, R, of such a clock may be expressed in the quadratic form

$$R = a + bt + ct^2 \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where a depends on the initial adjustment; b, the linear variation of rate, can usually be regarded as constant for periods of a year or so; and c is a small acceleration term which can often be neglected. The rate is given in milliseconds per day, and the change in the value is conveniently given in milliseconds per day per month; t is then the time in months. If the clock is used principally as a frequency standard, it is more convenient to express the values as parts in 10^9 and the variations as parts in 10^9 per month.

An observatory time standard usually consists of a group of clocks continuously intercompared. The comparison gives an indication of any departures of particular clocks from a uniform behaviour and enables the best clocks to be chosen for time prediction. A study of the clock performances has revealed a periodic variation in the rate of rotation of the earth.[1] It is now generally agreed that the amplitude of the seasonal variation found at first was too large,[2] and a reduced correction, ΔSV, given by

$$\Delta SV = 22 \sin 2\pi t - 17 \cos 2\pi t - 7 \sin 4\pi t + 6 \cos 4\pi t \quad . \quad (2)$$

where ΔSV is in milliseconds and t is the fraction of a year, has now been internationally accepted and is used to correct the time obtained by astronomical observations. A further small correction is made to allow for the movement of the poles, and the unit then obtained is known as the second of Universal Time 2 (U.T.2). The total amplitude of the time variation is about ± 30 millisec and that of the variation in rate is $\pm 0{\cdot}7$ millisec per day (± 8 parts in 10^9).

The paper is an official communication from the National Physical Laboratory.

Because of the seasonal variation and other irregular variations in the rotation of the earth on its axis, an alternative unit has now been defined by astronomers and adopted by the International Committee of Weights and Measures. This is based on the orbital motion of the earth and is defined as 1/31 556 925·974 7 of the tropical year 1900, January 0, at 12 h Ephemeris Time. This unit, known as the second of E.T. or simply as the second, is now the fundamental unit of time. Unfortunately, the astronomical measurements required to determine E.T. are considerably less precise than those required to determine U.T.2,[3] and they must be averaged over long intervals to give the required accuracy. The introduction of E.T. does not therefore help in the problem of calibrating and checking quartz oscillators.

The determination of a unit of time by means of an atomic standard is in marked contrast to the astronomical method and gives an accuracy of ± 1 part in 10^{10} in an interval of about 10 min. The performance of clocks can therefore be examined in detail. This study gives the maximum time interval which may be tolerated between calibrations by the atomic standard for a required accuracy and may also reveal the causes of the variations in the clock rates.

At the National Physical Laboratory the quartz clocks are calibrated in terms of both the caesium standard and, through time signals, the astronomical time scale U.T.2. The two calibrations give the relationship between atomic and astronomical time and the results obtained are summarized in the paper.

(2) SCOPE OF THE MEASUREMENTS

There have been three clocks in operation at the N.P.L. throughout the period of the tests (June, 1955–May, 1958) and a fourth was put into service in April, 1957. One of these clocks is calibrated directly in the manner described by Essen and Parry,[4] and the frequencies of the four are continuously compared by means of the monitoring equipment described by Steele.[5] A frequency of 2·5 kc/s derived from one of the clocks is conveyed by land-line to the Post Office Research and Development Laboratory (Radio Branch) at Dollis Hill and from there to the Royal Greenwich Observatory at Herstmonceux. It is compared at these laboratories with the local standards, which can thus be related to the caesium standard.

For the past three years the N.P.L. and the U.S. Naval Observatory, Washington, D.C., have co-operated in a programme of time-signal comparisons, which has enabled the unit of the time scale U.T.2, determined at the Observatory, to be accurately related to the caesium standard. The rates of the clocks are published in the Bulletins of the Observatory in terms of U.T.2 and can therefore be given in terms of the caesium standard.

The detailed analysis is restricted to the four N.P.L. clocks, and the monthly average values of the others are given as additional evidence of the general performance of the clocks. They all consist of a ring of quartz (Essen ring), manufactured, mounted and sealed in a vacuum enclosure at the Post Office Laboratory. In most of the installations the Post Office oven and circuit are used, but in others these components have been modified. The Post Office standard 19 differs also in being used as a resonator and not as an oscillator, and any ageing due to changes in the driving circuit is therefore eliminated. This standard is, moreover, sunk in a deep hole, where the ambient temperature is sufficiently constant to make further temperature control unnecessary. The clocks are listed in Table 1, together with the values of a and b [see eqn. (1)] as determined with reference to the caesium standard.

Table 1

Designation	Location	a	b
		$\times 10^{-9}$	$\times 10^{-9}$ per month
Q26	N.P.L.	+87·8	+2·56
Q13	N.P.L.	−64·4	+2·65
Q32	N.P.L.	−40·7	+4·71
Q60	N.P.L.	−4 138·9	+1·83
EB	P.O.	+698·5	+1·30
EA	P.O.	+902·5	+1·78
19	P.O.	+572·2	+0·70
E56	U.S.N.O.	−105	+4·50
E43	U.S.N.O.	+120	+4·62
C6	R.G.O.	+572	+2·93
E5	R.G.O.	−461	+1·12

(3) CALIBRATION IN TERMS OF THE CAESIUM STANDARD

The calibrations, described in detail elsewhere, are normal made at weekly intervals; their standard deviation is ± 1 pa in 10^{10}, and the time taken is about 10 min. A large number tests involving both the intercomparison of a number of cloc and their direct calibration in terms of the caesium resonan have established that during such intervals the clocks do no in general, depart from a value predicted by their immedia past performance by more than ± 2 parts in 10^{10}, and if a calibration point departs from a line representing a unifor behaviour by a greater amount the operation of the oscillat

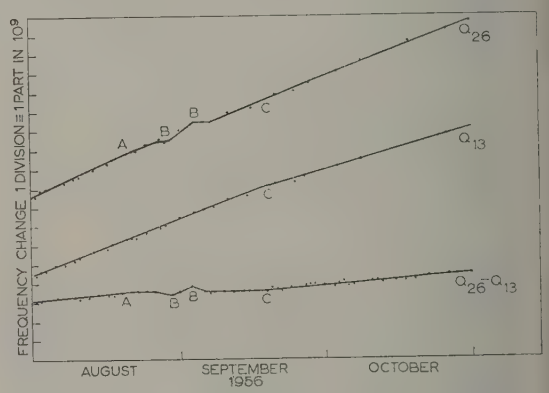


Fig. 1.—Absolute and relative changes in oscillators Q13 and Q26.

is investigated. The typical set of results reproduced in Fig. is selected because the clocks suffered several small chang during this period. The clock calibrated is Q26, and the resul for Q13 are obtained from the comparison results Q13–Q2 shown in the bottom curve. At A there was a very slig decrease in the drift rate of Q26 and at B there wa an irregularity in its performance which is clearly marke although it only persisted for a few days and amounted to $\pm$ parts in 10^{10}. At C there is a change in the drift rate of Q of about -1 part in 10^9 per month and possibly a very sma change in that of Q26. Further examples of abrupt changes rate are shown in Fig. 2, which gives the relative drift of clock Q13, Q26, Q32 and Q60. Sometimes, as at D, it is fairly eas to determine which clock has changed, but at others, e.g. E, F, and G, several of the clocks seem to have been affected different extents. It is generally possible, however, to obta the frequency of each clock in terms of the caesium frequenc with a precision of ± 2 parts in 10^{10}; and the results in Fig.

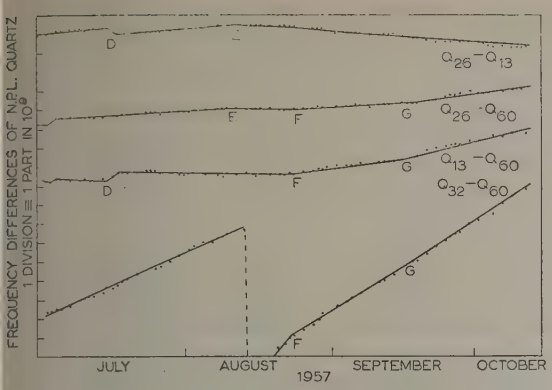


Fig. 2.—Variation in relative frequency drift of N.P.L. quartz oscillators.

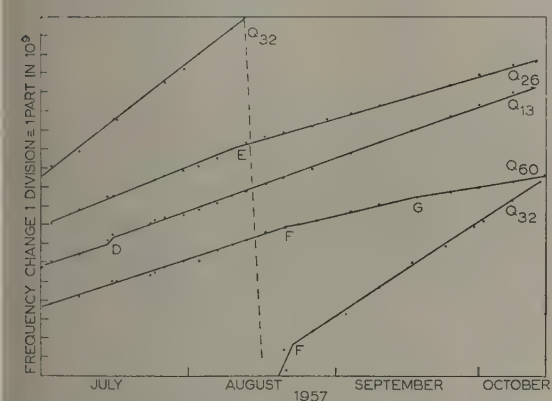


Fig. 3.—Variation in frequency drift of N.P.L. quartz standards in terms of caesium resonator.

e given in this form in Fig. 3. The discontinuity in the results Q32, which was due to an oven failure, is exaggerated to oid overlapping of the curves.

(4) LONG-TERM PERFORMANCE OF THE CLOCKS

The long-term performance of the clocks is displayed most adily if the linear drift term, *b*, is removed. The residual riations are shown in Fig. 4. The relatively larger scatter of e results for the clocks E56, E43, C6 and E5 is due to the ct that they are derived from values initially given in terms U.T.2. The errors in the derivation of the time scale and in e relationship between U.T.2 and the caesium frequency are erefore included. The results for C6 and E5 are given only the middle of 1957, because the Time Department of the Royal reenwich Observatory was then transferred from Abinger to erstmonceux.

The first point to notice is that the maximum variation of frequency from that given by a linear drift is ± 5 parts in 10^9, en for intervals as long as three years, and is only ± 1 part 10^9 for some intervals of one year. The results for several ocks such as Q60, E43 and possibly Q26 would be more uniform a small square term *c* were calculated and removed, but in ost cases the residual changes from linearity are not systematic. The irregularity of the N.P.L. clocks at the end of 1955, the ginning of 1956, and again in 1958 may be due to failures in

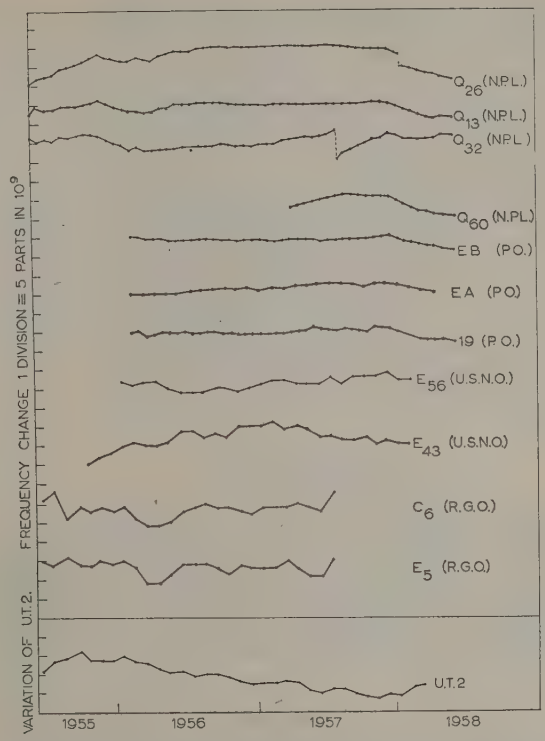


Fig. 4.—Residual frequency changes in quartz oscillators, after subtracting linear ageing rates: variation in U.T.2 expressed as a frequency.

the mains supply which were followed by a delay in the operation of the standby equipment. Such failures occurred in November, 1955, April and July, 1956, December, 1957, and January and March, 1958. During February, 1956, the voltage of the mains supply was so low that it fell outside the range for which automatic compensation is provided. Most of the failures were due to accidents and reorganizations internal to the Laboratory and not to interruptions of the public supply. They make it difficult to assess the performance of the clocks under good conditions, but perhaps it is some consolation to find that the performance was not upset more by the disturbances. The failure in January, 1958, produced an abrupt change in the frequency of Q26, but the more usual effect seems to be a change in the drift rate.

Although no systematic measurements have yet been made to determine the causes of the variations in rate, it is established that any interruptions in their operation, and particularly a failure in the temperature control, cause a deterioration of performance for some time afterwards.

(5) VARIATIONS IN THE UNIT OF U.T.2

In 1955 a joint programme was arranged between the N.P.L. and the U.S. Naval Observatory with the object of determining the frequency of the caesium resonance in terms of the second of E.T.,[6] and this programme has also provided accurate information concerning the variations of U.T.2.[7]

The intervals between specified time signals from the stations GBR (Rugby) and WWV (Washington) were measured at the N.P.L. by reference to one of the quartz clocks calibrated by the caesium standard, and at the U.S. Naval Observatory by reference to the similar quartz clocks calibrated by astronomical

observations. In this way the actual errors of the signals were eliminated and it was estimated that the errors due to radio propagation and to instrumental effects did not exceed 1 millisec. The monthly average values were thus accurate to ± 3 parts in 10^{10}. The information was exchanged in such a way as to ensure that the two measurements were quite independent.

The change in the unit of U.T.2 compared with the caesium frequency is given as the bottom curve in Fig. 4. For a period of two years there was a steady slowing down of the rate of rotation of the earth of about 5 parts in 10^9 per year. If the corrections for the seasonal variation [eqn. (2)] and for the movement of the poles are put back into the values of U.T.2 the variation of the unit of Universal Time U.T.0 is obtained. This is shown in Fig. 5.

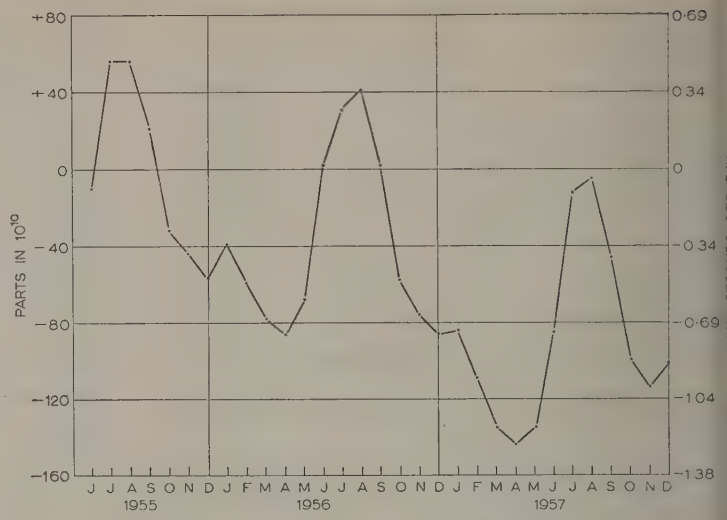


Fig. 5.—Variation in the rate of rotation of the earth.

(6) ACKNOWLEDGMENTS

Acknowledgment is made to the Astronomer Royal, the Engineer-in-Chief of the Post Office, and the Superintendent of the U.S. Naval Observatory for permission to use the results for their quartz standards. The work was carried out as part of the research programme of the National Physical Laboratory and is published by permission of the Director.

(7) REFERENCES

(1) SMITH, H. M.: 'The Determination of Time and Frequency', *Proceedings I.E.E.*, Paper No. 1054 M, November, 1950 (**98**, Part II, p. 143).

(2) SMITH, H. M., and TUCKER, R. H.: 'The Annual Fluctuation in the Rate of Rotation of the Earth', *Monthly Notices of the Royal Astronomical Society*, 1953, **113**, p. 251.

(3) MARKOWITZ, W.: 'Photographic Determination of the Moon's Position and Applications to the Measure of Time, Rotation of the Earth and Geodesy', *Astronomic Journal*, 1954, **59**, p. 69.

(4) ESSEN, L., and PARRY, J. V. L.: 'The Caesium Resonator a Standard of Frequency and Time', *Transactions of t Royal Society*, 1957, **250**, p. 45.

(5) STEELE, J. McA.: 'The Standard Frequency Monitor at t National Physical Laboratory', *Proceedings I.E.E.*, Pap No. 1765 M, October, 1954 (**102**, Part B, p. 155).

(6) MARKOWITZ, W., HALL, R. G., ESSEN, L., and PARR J. V. L.: 'Frequency of the Caesium Resonance in Ter of Ephemeris Time', *Physical Review Letters*, 1958, p. 105.

(7) ESSEN, L., PARRY, J. V. L., MARKOWITZ, W., and HA R. G.: 'Variation in the Speed of Rotation of the Eart *Nature*, 1958, **181**, p. 1054.

DISCUSSION BEFORE A JOINT MEETING OF THE MEASUREMENT AND CONTROL SECTION AND THE ELECTRONICS AND COMMUNICATIONS SECTION, 1ST DECEMBER, 1959

Mr. R. L. Corke: The time-signal and standard-frequency transmissions sent out from Rugby by the Post Office, with the help of the Royal Greenwich Observatory and the National Physical Laboratory, are made with respect to the two time scales discussed in the paper, and this is a matter of difficulty which I will explain. Time signals are related to the U.T.2 scale by the Observatory and thus to earth time by the *Bulletins* issued from time to time by the Observatory. Frequency is measured against the E.T. scale, which is represented by the resonance of caesium. Our standard-frequency transmitters at Rugby are controlled entirely from one master oscillator, and during 1959 its frequency has been offset by -170 parts in 10^{10} so that the time signals controlled by the same oscillator do not drift too quickly from U.T.2, otherwise frequent phase adjustment would be necessary to keep within the 50 millisec tolerance allowed. The mean value for U.T.2 adopted for 1960 is -150 parts in 10^{10}. Disturbances to oscillators to reset their frequency in this way are not conducive to stability and should be avoided. Many observations on oscillators have shown that there is no systematic behaviour which enables us to forecast the performance of a particular oscillator after a disturbance, and months sometimes pass before an oscillator has settled after a severe disturbance such as loss of temperature control. We could avoid most tuning adjustments by using continuous phase shifters, and these cou be used to adjust independently the effective frequencies of t standard-frequency control and the time-signal generator. W could thus satisfy the N.P.L. and R.G.O. without the need f annual retuning according to the trend of U.T.2.

At the Post Office Research Station we are experimenting wi a caesium resonator based on the N.P.L. design, and we a grateful to the authors for their help in this work. We ha become aware during the progress of these experiments that t caesium resonator in its present form is liable to many sm errors, and it requires most skilful experimental procedure to u it properly. A recent paper* examines the difference betwe measurements of the frequency of the caesium resonance he and in the United States: among the possible causes of error a non-uniformity of the magnetic field in the interaction spa lack of phase synchronization and difference in relative directio of the magnetic a.c. fields in the two Ramsey resonators, t presence of sidebands in the search oscillator and transfer erro between this and the oscillator being calibrated. It is o experience that oscillators which do not have a reasonably goo

* HOLLOWAY, J., MAINBERGER, W., REDER, F. H., WINKLER, G. M. R., ESSEN and PARRY, J. V. L.: 'Comparison and Evaluation of Caesium Atomic Beam F quency Standards', *Proceedings of the Institute of Radio Engineers*, 1959, **47**, p. 173

ng-term frequency performance will not have a good short-rm performance and are therefore difficult to calibrate. During e earlier part of the period of review, when the N.P.L. oscil-tors were not very stable, there must have been occasions when equency calibration was difficult and errors may have arisen. roups of oscillators—probably some of those shown in Fig. 4—st revealed variations of the speed of rotation of the earth. he same oscillators appear to show, in December, 1957, a riation in the apparent frequency of calibration by caesium of rtain oscillators, and some measurements of the frequencies of deeply buried resonator and oscillator EB made on the installa-on at Dollis Hill show the change rather clearly.

I agree that quartz oscillators are likely to continue in use as ocks for a long time. Until we have a caesium-controlled scillator running continuously as a clock, it is, I think, a isnomer to call the caesium resonator a clock. Reliance on a ngle caesium resonator was necessary in the early period of its troduction, but when caesium devices became available in the nited States it was possible to intercompare them by calibra-ons of radio transmissions. These have been mainly one-way rough the medium of MSF (60 kc/s) and GBR (16 kc/s). But n the 1st December, 1959, the U.S. Navy transmitter NBA 8 kc/s) in the Panama Canal Zone reopened with its carrier alibrated against caesium and its time signals calibrated by the .S. Naval Observatory, Washington, thus giving for the first me a good oversea v.l.f. signal in the United Kingdom for ccurate comparisons. Although experimental as yet, the trans-issions from NBA will enable continuous intercomparisons to e made between the American and British standards of frequency nd time and will improve the world-wide service.

Mr. L. A. Thomas: The precision with which the frequency f the caesium atomic resonance is quoted prompts me to ask ow much the observed frequency is affected by the microwave chnology. The authors have shown the response curves for vo cavity systems of which the second has a much smaller ne width. What bearing has the line width on the precision of ne caesium standard and how has the reduction in line width een brought about?

To put the caesium standard in perspective it may help to nsider the degree of frequency control required for present-day plications in radiocommunication. For operation over a ide temperature range a stability of a few parts in 10^5 may be lerated, but more typically the requirement is for a few parts 10^6 in mobile equipment and in 10^7 for fixed stations. Beyond is there is a trend towards still tighter tolerances on the fre-uency of transmitted signals corresponding to stabilities of the rder of parts in 10^9 or even 10^{10}. As a step towards meeting ese more stringent requirements, new types of quartz crystal e now under development, aimed at a stability of a few parts in 0^8. In order to monitor the long-term tests inevitably associated ith their development, it is necessary to have access to standards ith a precision of at least several parts in 10^9. In this connec-on, effective access to a high-precision signal based on the esium standard is of considerable importance.

The authors' observations on variations in the rotation of the rth are only a sample of the possibilities which the caesium andard may offer in the fields of geophysics and astrophysics, cluding the topical subject of interplanetary navigation. At ne same time consideration must be given to the feasibility of her frequency-control systems depending on atomic or molecu-r resonances.

Mr. F. G. Clifford: Fig. A shows the intercomparisons of the equencies of two units of the Post Office standard and that of the .P.L. caesium resonator in a period of nearly four years. Both nits are controlled by a ring-type crystal, EB being a Meacham ridge oscillator and No. 19 being a resonator sited at the bottom

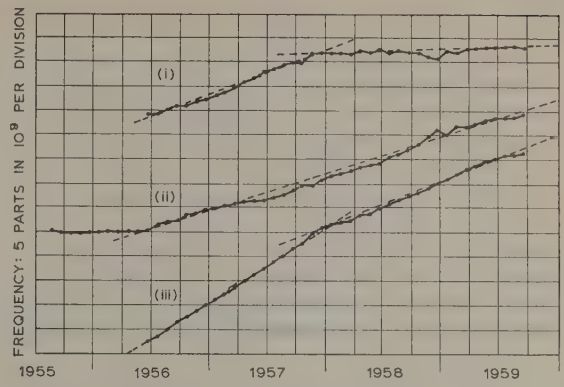


Fig. A.—Frequency comparison.
(i) 19-caesium standard.
(ii) EB-19.
(iii) EB-caesium standard.

of a 60 ft bore hole. The variation of EB with No. 19 lies substantially on a straight line; variations of EB and No. 19 with the caesium standard are also substantially linear, but changes in the rates of about 6 parts in 10^{10} per month occurred in December, 1957. These changes are not fully understood, but similar changes are shown when the N.P.L. caesium resonator is compared against other oscillators, e.g. the MSF units at Rugby and the N.P.L. No. 13.

The good performance of No. 19 is no doubt largely due to the constant and low temperature (11·5° C) and freedom from mechanical shocks. Resonators, however, are inconvenient to use, and the Post Office has under test a transistorized oscillator, using a ring-type crystal, at the bottom of a bore hole. The results are very promising. To overcome the difficulties of making frequency adjustments, and the resulting unpredictable variations in frequency, I have proposed that continuous phase shifters should be used in conjunction with the oscillator. One phase shifter could control the frequency and another the time signal transmissions, and it would be no longer necessary for the frequency offset between the MSF Rugby transmissions and the caesium standard.

Until such time as an atomic frequency standard which is truly definitive and has a long life is produced, I favour the idea of a frequency-standard installation including several standard oscillators, some of them buried, some buried resonators and at least two caesium resonators.

Miss C. J. A. Penny: By courtesy of the Director of the National Physical Laboratory, the caesium figures are made available to the Royal Greenwich Observatory, and these, together with a direct line between the Observatory and the Laboratory, give us most of the advantages of having a caesium standard without the work attached to it.

It was a little unfortunate that the two Observatory clocks shown in Fig. 4 were compared via the radio link. Although the direct link was not in operation until 1957, there was an experimental link between Abinger and the Post Office and this, together with another between the Post Office and the N.P.L., has enabled us to make a better comparison with the caesium standard.

The Abinger clocks are shown at the top of Fig. B, and the scale is more than twice that shown by the authors. Underneath are shown two clocks now in operation at Herstmonceux, covering the period 1957–59. The divergence, which has already been mentioned, between the clocks and the caesium standard at

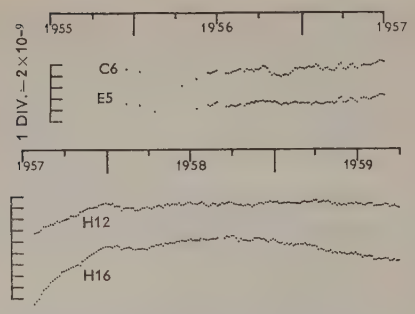


Fig. B.—Comparison of Royal Observatory quartz clocks and N.P.L. caesium standard.

the end of 1957 is evident. There is probably some residual ageing in H16, but H12 had been running continuously since the beginning of the year. All our other clocks showed this divergence at this time.

During the period 1955–59 there were a number of disturbances to the clocks and astronomical instruments occasioned by the move of the Observatory, and it was fortunate that it has been possible to make comparisons with United Stat observations.

The frequency variation of U.T.2 as determined at Greenwic and Herstmonceux agrees exactly with the authors' figure f 1955–57, i.e. the rate of retardation of the Earth is 5 parts in 1(per year. After that time it would appear that the U.T.2 tim system has suffered no secular change with respect to the caesiu standard.

Mr. C. S. Brown: Will the authors comment on the use c servo-corrected quartz oscillators as used in the America Atomichron, and also state whether they think that the caesiu beam standard has reached the stage where we can consider it a mobile device?

From experiments which are being carried out at the Nation Bureau of Standards Laboratory on the use of quartz oscillato at liquid-helium temperatures, Q-factors as high as 4×10^7 hav been quoted. Bearing in mind the improved ageing obtaine with the buried resonator operated by the Post Office, where th temperature is only slightly lower than the normal operatin temperature, can we expect considerable improvements in th ageing rates of quartz clocks in the future, together with bette short-term stability?

Mr. L. Stenning: Was the buried resonator No. 19 left quiescen between measurements or kept continuously excited?

THE AUTHORS' REPLY TO THE ABOVE DISCUSSION

Dr. L. Essen, Mr. J. V. L. Parry and **Mr. J. McA. Steele** (*in reply*): Some reference has been made to the N.P.L. caesium resonator, which, while it is not considered explicitly, does, of course, form the basic standard of comparison for all the results given in the paper. Rather unhappily, it seems to us, Mr. Corke has stressed the difficulties and uncertainties associated with the operation and realization of a caesium standard, for our experience points in the opposite direction. The original beam assembly at the N.P.L. incorporated many facilities for varying parameters thought likely to affect the frequency of the central resonance, and after many careful experiments with this equipment we were able to conclude that the caesium resonator can be used to define frequency with a standard deviation not exceeding 1 part in 10^{10}. This figure includes both systematic and random effects and is obtained regularly in the routine operation of the standard.

In reply to Mr. Thomas, the line width of the resonance, Δf, is given by $\Delta f = 0{\cdot}65\alpha/l$, where α is the most probable velocity of the atoms and l is the distance between the cavities in the Ramsay method of separated excitation, in which the beam traverses an r.f. field twice in its passage to the detector. The relative phase of the fields in the two cavities must be controlled to about $180/\Delta f$ deg if errors greater than 1 part in 10^{10} are to be avoided.

Mr. Corke, Mr. Clifford and Miss Penny have all referred to or elaborated on the change in frequency of a number of oscillators at the end of 1957. Around this time there is no internal evidence which would lead us to doubt the accuracy of the caesium calibration or the means by which it was transferred to oscillators in other laboratories. It should be noted, however, that frequency comparisons using both long- and short-wave signals had shown a divergence amounting to one or two parts in 10^9 between November, 1957, and January, 1958, between the N.P.L. standard and the Atomichrons in the United States. It was to resolve this discrepancy and at the same time eliminate the possible errors in radio transmission that two Atomichron were later sent to the N.P.L. for direct frequency comparisons and it is satisfactory to record that the frequency difference between one or other of these units and the N.P.L. resonato have not exceeded 4 parts in 10^{10} from March, 1958, to th present day. If the frequency of the N.P.L. caesium standar were increased by 2 parts in 10^9 for the months at the end of 195 the clock results would look somewhat smoother but the vari ation in the rate of rotation of the earth in the opposite sens would be more abrupt. It is clear that no alteration in a singl variable can be reconciled with the wide variation in the chang of drift rate over this period, extending from practically zero fo E56 to more than 1 part in 10^9 per month for H16—one of th oscillators cited by Miss Penny—and a change in the opposit sense in the rate of rotation of the earth.

The question of servo controls was raised by Messrs. Brow and Clifford. We have not yet experimented with a servo loo between the caesium standard and quartz oscillator, and althoug a servo control would be convenient, we doubt whether it woul at present improve the accuracy obtained. A quartz oscillato calibrated by the caesium standard might be a better clock tha one controlled by the standard. We think that atomic equip ment may be regarded as being mobile in the sense that it ca be flown across the Atlantic and operates satisfactorily as soo as it is connected to the appropriate power supply.

At the N.P.L. we have obtained Q-factors of about 10 from ring crystals operated in the range 2–4° K, and althoug long-term experiments are still in train to verify the point, w expect the normal ageing processes to be completely absent We may not, however, be able to exploit fully the increase stability, owing to the much enhanced shock sensitivity o quartz crystals at these temperatures.

Regarding Mr. Stenning's inquiry, we understand from th Post Office that the buried resonator 19 is excited only at th times of measurement.

:1.3.011 : 389.6(100)

The Institution of Electrical Engineers
Paper No. 3242
May 1960

ELECTRICAL UNITS AND STANDARDS

A Review of Progress

By P. VIGOUREUX, D.Sc.(Eng.), Associate Member.

(1) INTRODUCTION

A progress review[1] published in 1952 described the functions the international organizations which concern themselves with ectrical units and standards; it recorded the then recent change ade to the numerical values of material standards to bring em into agreement with the 'absolute' units, i.e. with electrical its based on the internationally agreed definition of the npere in terms of the force between straight parallel filaments which electric currents are maintained; it dealt in some detail ith the rationalized M.K.S., or Giorgi, system of units; and it scussed the interpretation of 'quantity equations'. With one inor exception all that was said then still holds good and it is nnecessary to repeat any of it. The exception is the definition the second of time, which was envisaged in the 1952 review id is dealt with in Section 2.1 of this one.

The position with experimental work is different in that much iat is novel has been done in the interval. The development lls naturally into three classes, two of which were mentioned the 1952 review as likely to have an important bearing on the ture position of electrical units and standards: atomic beams ave been used, notably at the National Physical Laboratory, ngland (N.P.L.), to measure frequency far more accurately ian has ever been done before; a physical constant, the gyro-agnetic ratio of the proton, has been measured at the National ureau of Standards, U.S.A. (N.B.S.), to a greater accuracy than as been done before; and, finally, a remarkable investigation y the National Standards Laboratory, Australia (N.S.L.), has rovided an alternative to the hitherto universal use of inductance s the fundamental standard from which the unit of resistance derived for accurate work.

It seems that we are now witnessing a change in the position f electric units and standards the like of which has not been en since F. E. Smith and his collaborators some fifty years ago id the foundations of accurate determination and maintenance f units by their work on the current balance, the Lorenz pparatus, standard cells, voltameters and mercury resistors.

(2) FREQUENCY

(2.1) Definition of the Second of Time

In the course of the last eight years, quartz crystal oscillators, reviously mentioned[1,2] as forming material standards more ccurate than any other electrical standard, have continued to be sed for linking astronomical observations, and their reliability as if anything been improved. They are mentioned again ecause to a great extent it is this very reliability and constancy ver periods of weeks and even months that made possible ie remarkable success achieved at the N.P.L. with the caesium omic beam as a time-keeper.[3]

These same oscillators used as intermediaries between the omic beam and astronomy have made an accurate definition f the second possible. As, however, they have also shown that ie rate of rotation of the earth varies periodically throughout ie year,[3,4] and as a study of the solar system indicates that the period of rotation of the earth has changed in an irregular way by as much as 80 parts in 10^9 in the past two centuries,[5] it has been necessary to base the definition of the second on the period of revolution of the earth about the sun. The period chosen has been that of the 'tropical' year 1900. Time so defined is called Ephemeris Time, and in 1956 the International Committee of Weights and Measures[6] decided that 'the second is the fraction 1/31 556 925·974 7 of the tropical year for 1900, January 0, at 12 hours Ephemeris Time'.

(2.2) Caesium Beam

Atoms of one and the same element can exist in a number of different states corresponding to different configurations of the electrons surrounding the nuclei. Each state is characterized by a particular energy, and changes from one state to another are accompanied by emission or absorption of electromagnetic radiation of frequency, f, given by

$$f = \frac{W}{h} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where W is the difference of the energies of the two states and h is Planck's constant.

On account of the spins with which they are endowed atoms act in an external magnetic field like small magnets, and for caesium atoms two energy states, designated by F(4, 0) and F(3, 0), differ in that the atomic dipole moments are opposite in direction for the two. The difference of energy between the states corresponds to a frequency difference of about 9·2 Gc/s, which falls conveniently in the radar centimetre range and is almost independent of the magnetic field. Transitions from one state to the other are stimulated if the atoms in the constant magnetic field, H, are exposed to an electromagnetic field of frequency satisfying eqn. (1).

Essen and Parry[7,8] made use of these properties by employing Rabi's double-deflection method, shown in Fig. 1. Atoms streaming from the oven into a highly evacuated space enter a strong magnetic field, H_1, which has a strong gradient in a

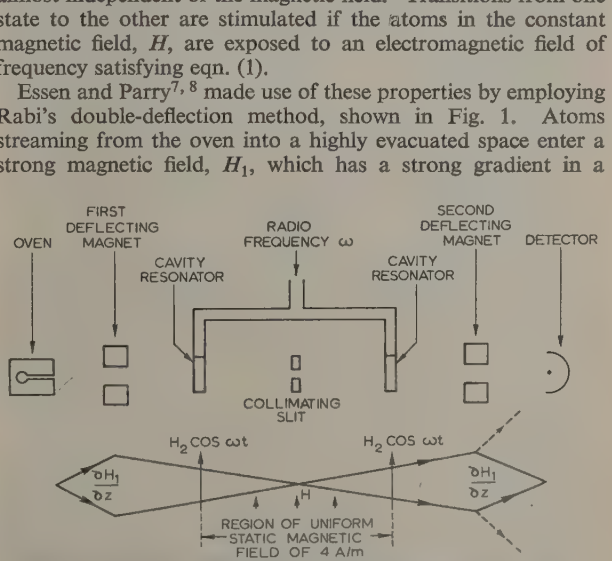


Fig. 1.—Principle of detection of resonance of caesium atomic beam.

Dr. Vigoureux is at the National Physical Laboratory.

direction perpendicular to their path. Those of one state are deflected one way, those of the other the other way by angles dependent on their velocities; they then enter a region of uniform magnetic field, H, where a number come to a focus in the central slit, go through it and travel past the region of uniform field, H, into the gap of a second magnet, with a field gradient similar to that of the first magnet. If there have been no transitions the atoms are deflected further out, as shown by the dotted lines, and have no further effect; but those which have undergone a transition in the space between the two magnets are deflected an equal amount the opposite way and come to a focus on the detector. Transitions are caused by radio-frequency fields, H_2, applied to the two cavity resonators and are more or less numerous according as the frequency of the field is exactly as given by eqn. (1) or not. The detector consists of a tungsten strip heated to about 1 000° C, at which temperature the caesium atoms striking it are emitted again as positive ions and are collected by a curved plate maintained at a potential of −20 volts relative to the wire. The collector plate is connected through a sealed lead to a vibrating-reed electrometer, in which the voltage developed across a 10^{10}-ohm resistor is converted to an alternating voltage and amplified. The amplified signal is rectified and a reading proportional to the beam intensity is obtained on a microammeter.

The procedure is to vary the radio frequency by small amounts and to plot a resonance curve from which the central frequency is determined. This frequency is next compared with one derived from quartz oscillators by methods involving multiplication and other processes which it would take too long to describe here. Suffice it to say that the resonance curve is only 330 c/s wide at the half-power points and that measurements taking about 10 min provide a setting accuracy within 1 part in 10^{10}. The width of the resonance curve is inversely proportional to the distance between the cavity resonators: the width given here is for a separation of about 50 cm.

Since the uniform magnetic field, H, in which transitions occur affects the frequency, if only very slightly, it is made small. As, however, it also splits each of the two energy states into a number of 'Zeemann' levels separated by small intervals proportional to it, it must be made large enough to eliminate interference by resonances between subsidiary levels of the two main states. The value selected for the N.P.L. beam is 4 A/m, which is large enough to resolve Zeemann lines although it changes the frequency derived from the central lines by less than 1 part in 10^{10}.

A three-year collaboration between the U.S. Naval Observatory for the astronomical observations and the N.P.L. for the provision of time based on the frequency of the caesium-beam resonance has resulted in the latter being determined as 9 192 631 770 ± 20 c/s of Ephemeris Time.[9]

The N.P.L. caesium beam was the first to work in a manner satisfactory for time measurement. Other caesium beam standards have now been made, notably in the United States.[10] One has also been made by the National Research Council of Canada (N.R.C.).[11] The 'atomichrons' described in Reference 10 differ from the N.P.L. standard in a number of ways, but especially in that they contain an electromechanical servo system which maintains the frequency of an oscillator at the caesium beam frequency whereas the N.P.L. standard is used as a resonator. It was therefore of great interest to compare the two types. This comparison has been carried out, by radio in 1957[12] and at the N.P.L. in 1958.[13] The two complete atomichrons and an experimental tube differed from the N.P.L. standard by between 3 and 5 parts in 10^{10}, but when they were used as resonators these differences were reduced by 2 parts in 10^{10}. Until comparisons with other standards are made it may be assumed that the caesium beam resonance is reproducible by standards of different type to about 2 parts in 10^{10}.

Although quartz crystal oscillators are still a very convenient intermediary between the caesium beam resonance frequency and astronomical observations, the success achieved with the atomichron has shown that systems can be developed to produce a signal at the caesium beam frequency for periods long enough for a direct link with astronomy. Thus quartz oscillators may cease to be essential for accurate time keeping. Indeed they may soon become superfluous, for it is conceivable that other atomic standards of frequency will be developed, e.g. using 'optical pumping'[14, 15] of rubidium vapour, which might prove more convenient and easier to use than the caesium beam. In this method the vapour is irradiated by monochromatic light which causes transitions from a low level A to a level C from which the atoms fall either back into A or into a state B separated from A by a frequency in the microwave spectrum. The process therefore results in more and more atoms being transferred or 'pumped' into level B, from which equilibrium between A and B is restored by a radio-frequency field of frequency corresponding to the transitions A ↔ B.

(2.3) The 'Second' of the Future

In 1958 the International Committee of Weights and Measures[16] agreed to submit to the eleventh General Conference of Weights and Measures in 1960 a recommendation to define the metre as equal to 1 650 763·73 vacuum wavelengths of the radiation of the orange line of krypton-86. There is little doubt that the Conference will ratify this recommendation and that the metre will no longer be defined as the length between lines engraved on a platinum-iridium bar. The new definition has among others the advantage that the frequencies emitted by atoms are far more permanent and reproducible than even the best material standards.

This impending change immediately brings to mind the question of the definition of the second. Astronomical observations, even averaged over several years, do not yield anything like the accuracy with which frequencies can be compared in ten minutes, and the solar and other heavenly systems are subject to secular changes which, as far as we know, atomic processes do not suffer. It seems, therefore, that there is much to say for doing with the second what is being done with the metre.

If, moreover, it is borne in mind that caesium standards, even without the rather strict specifications which have been found desirable for the new length standard, can be reproduced to at least ten times the accuracy with which Ephemeris Time can be measured even over many years, it seems that there is some prospect[5, 17–20] of defining the second of time in terms of the frequency of an atomic radiation with a reproducibility of 1 part in 10^{10}.

(2.4) Radio Transmissions of Standard Frequencies

The MSF transmissions[21] started experimentally in 1950 are now operated from the Post Office station at Rugby on behalf of the N.P.L.; they are intended to serve mainly the United Kingdom and Western Europe, and the reports received over the last five years indicate that their object is being achieved.

Continuous transmissions are made on 2·5, 5 and 10 Mc/s, and for one hour daily on 60 kc/s, and the carriers are modulated by pulses at 1 c/s according to a published programme. These transmissions have proved so useful that the carrier wave of the more powerful 16 kc/s GBR telegraph transmitter is now also controlled by the MSF standard, although it is not part of the MSF service. Standards in the United States and New Zealand have been compared with this carrier to an accuracy exceeding 1 part in 10^9.

Since June, 1955, the frequencies have been measured in terms of the caesium atomic frequency standard of the N.P.L. and have been maintained within 5 parts in 10^9 of nominal value, i.e. of the value in terms of the 'second of uniform provisional time', also denoted abroad by TU2 or UT2; this unit of time is based on the period of rotation of the earth with the seasonal variation smoothed out.

Since March, 1959,[22] in view of the official adoption of the second of Ephemeris Time as the unit of time by the International Committee of Weights and Measures,[6] the corrections have been based on a frequency of 9 192 631 770 c/s for the caesium standard and they have been given to within 1 part in 10^{10} instead of the previous 1 part in 10^9.

(3) ELECTRICAL UNITS

(3.1) International Comparisons

Every few years the International Bureau of Weights and Measures compares the units of resistance and of electromotive force maintained by national laboratories, who provide standard resistors and cells for the purpose. The last results for the ohm[23, 24] show that, with the exception of one laboratory which has subsequently adjusted its value to that of the N.B.S., the mean deviation of the units maintained by eight national laboratories is less than 3 parts in 10^6, and that the stability is of the order of 1 in 10^6 in two years. The figures for the volt[25, 26] are similar.

The countries concerned do not as a rule adjust their units as a result of these comparisons: it is better to maintain continuity as long as the differences are not serious.

(3.2) Current

The international system of electrical units, known also as the M.K.S., Giorgi or M.K.S.A. system, is based on the definition of the ampere in terms of the metre, kilogramme and second, but the ampere need not be the first electrical unit to be established absolutely. The definition can instead be applied indirectly to the determination of other units, as for instance inductance or resistance, and strictly speaking once one of the units, e.g. that of resistance, has been established, the other two principal units, namely those of current and potential, can be obtained from the laws of Joule and of Ohm. As, however, calorimetric methods are inconvenient, standardizing laboratories normally establish two of the units by electrical methods based on the definition of the ampere and on the accepted laws of electromagnetism. In general, the ampere is determined by some form of dynamometer almost directly from the definition, and the ohm is established indirectly from the definition and the law of electromagnetic induction.

The most recent determinations of the ampere are those by Driscoll[27] and by Driscoll and Cutkosky[28]. Two dynamometers were used, one a Pellat type which measures the torque between two concentric coils when their axes are at right angles, the other a Rayleigh-type current balance which measures the axial force between coaxial coils. The results of the two determinations differed by 5 parts in 10^6, and the mean result indicates that the unit of current maintained at the N.B.S. by standard resistors and cells is 10 parts in 10^6 greater than the ampere.

The N.B.S. has also recently completed[29, 30, 31] the determination of a physical constant which is closely related to the ampere as regards experimental method. The proton or nucleus of the hydrogen atom has magnetic moment, M, and angular momentum L; we do not know enough about the atom to calculate M, so that the quotient M/L, or γ, called the gyromagnetic ratio, must for the time being be regarded as an experimental constant of nature. One method of determining it is to align protons by application of a strong magnetic field to some liquid containing hydrogen, and allow them to regain thermal equilibrium in a known constant magnetic field H. On account of their angular momentum they do not straightaway realign themselves with the field but instead precess about it, thereby producing an alternating magnetic field which persists for some time if the source of protons is a liquid; this time is about 3 sec for water and about 16 sec for benzene; it can be shown[32] that the angular frequency, ω, of the field is γH. A method of determining γ is to measure the frequency of precession produced by a known constant field.

Bender and Driscoll[31] produced this field by maintaining in a solenoid of measurable dimensions a constant current known in terms of the ampere previously determined at the N.B.S. The water is contained in a glass sphere 20 mm in diameter magnetized in a separate building 15 m away by inserting it in the gap of a strong magnet; it is then shot through a pneumatic tube and comes to rest at the centre of the solenoid. In the process, which takes about a second, the magnetization does not decrease very much but it assumes the direction of the field of the solenoid. As, however, in order that precession may produce the greatest external alternating field, the magnetization must initially be at right angles to the direction of the constant field, it is first shifted 90° by application of a pulse at precession frequency, after which the protons are allowed to precess, so inducing an e.m.f. in a detecting coil surrounding the sphere. This coil is connected to an amplifier and counter for comparison of the frequency with the standards of the Bureau.

Final results of this ingenious and elegant work are not yet to hand, but the preliminary value[33] obtained for γ uncorrected for the diamagnetism of the water sphere is 336·167 m/As. The error, which seems to be a so-called '50% error' due partly to the uncertainties of the measurement described here and partly to those of the determination of the ampere mentioned above, is given as 8 parts in 10^6.

Some may, with good reason, look upon this remarkable investigation as the first sign of an impending change in the position of electrical units and standards. When current is measured by some form of balance, the force between the coils is obtained by weighing and thus entails a knowledge of the acceleration due to gravity: moreover, the force depends to a great extent on the coil diameters, which must therefore be accurately known. The measurement of the gyromagnetic ratio, on the other hand, is based on a calculation of the magnetic field at the centre of a coil, usually a long solenoid, and as is well known the field then depends chiefly on the turns per unit length and very little on the diameter. The other measurement needed is that of frequency, which in principle is more accurate than almost any other measurement. Thus it might well be argued that, granted time to overcome the many difficulties which still beset a determination of gyromagnetic ratio, the operations involved are inherently capable of an accuracy higher than that of a measurement of current by a current balance. It has even been estimated that γ should be measurable to an accuracy of 2 or 3 parts in 10^6.

Hitherto the national laboratories of the world have relied on their material standards, i.e. standard resistors and cells, to maintain the ampere and to compare absolute determinations, and as these standards are good to 2 or 3 parts in 10^6 whereas the extreme limits of uncertainty (not the probable error) of a measurement of current with one of the best current balances in the world[34] is assessed at some 20–30 parts in 10^6, there has been a comfortable margin. If, however, absolute determinations become more accurate than material standards, the latter cannot serve for comparison as now understood, but

only as a safeguard—still an eminently useful one—that the optimism of investigators shall not lead to assessments of accuracy ten times too high.

(3.3) Capacitance

Hitherto national laboratories have established their electrical units by maintaining, in addition to dynamometers to measure current, inductors which serve to determine resistance and capacitance.[35] The N.P.L., for instance,[36, 37] the N.R.C.[38, 39, 40] and the N.B.S. have mutual inductors of the Campbell type. In 1951, Rayner,[36] using the standard mutual inductor of the N.P.L., verified that the standard resistors of that laboratory had not changed by more than a few parts in 10^6 since the previous determinations in 1936, and he found that the unit maintained there by means of those resistors was $1{\cdot}000\,002 \pm 0{\cdot}000\,015$ ohms. Measurements of that sort, depending like those of current on a knowledge of the diameters and lengths of coils of wire, are subject to uncertainties of 10–20 parts in 10^6.

This position is also likely to be changed, this time by the work of Thompson, Clothier, Lampard and others at the N.S.L. The story starts by Lampard[41, 42] proving that the capacitance in vacuo between two segments like CB and AD of a right cylinder of any shape symmetrical about AB, Fig. 2, is equal

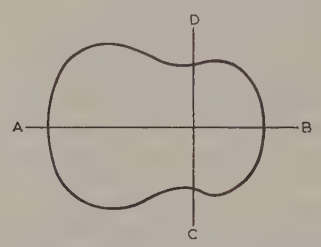


Fig. 2.—Cylindrical capacitor with cross-section symmetrical about AB.

to $\epsilon_0 (\log_\varepsilon 2)/\pi$ per unit length, where ϵ_0 is the permittivity of vacuum, and does not depend on the shape of the cross-section. In practice, there must be gaps between the four segments, since measurement of capacitance involves differences of potential, but Thompson, who had suggested the investigation of cylindrical cross-capacitors,[41, 43] proposed one made of four insulated rods (Fig. 3), with which, owing to the re-entrant shape, the configura-

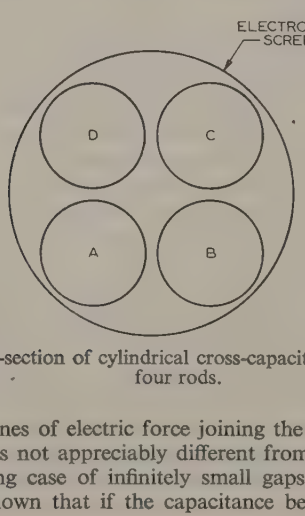


Fig. 3.—Cross-section of cylindrical cross-capacitor constructed from four rods.

tion of the lines of electric force joining the 'live' segment and its opposite is not appreciably different from what it would be in the limiting case of infinitely small gaps. Lampard[42] has, moreover, shown that if the capacitance between A and C is nearly equal to that between B and D all errors due to asymmet disappear from the mean value. In practice, equality can achieved to a high order of accuracy by fairly simple adjustme of the position of one of the rods, and the formula is th accurate to at least 1 part in 10^7.

In order that the formula may be applicable, end effects mu be eliminated. At the N.S.L., Clothier[44] eliminates them inserting axially at the two ends of the system, in the spa between the four rods, cylindrical shielding tubes electrical connected to the outer screen. If each cross-capacitance measured for two separations of the ends of the shields and t difference of the means taken, the formula can be applied exact to the change of separation. This change is measured direct in wavelengths of the green radiation from an electrodele mercury-198 isotope lamp with the aid of a Fabry–Perot inte ferometer mounted in the ends of the shield tubes. The capacit is erected with its axis of symmetry vertical, and the upper tu can be translated along that axis through a total travel of abo $5\frac{1}{2}$ in by means of a screw. The lower tube can be translat vertically through a range of a few wavelengths and can also tilted in two perpendicular directions for fine adjustment of t interferometer plates for parallelism. The adjustments to t lower shield tube are performed electromagnetically. In additio electromagnetic control is made to impart a small vertic oscillation to the lower shield tube and thus to the optical pa difference. A photomultiplier detects the corresponding cyc variations in the intensity of the central fringe, the mean intensi of which can be brought to a maximum by vertical translati of the lower shield. In use, the shields are varied in separati by an integral number of half wavelengths, the initial and fin positions being located at intensity maxima for the centr fringe. The actual number of half wavelengths in the interv is determined by a build-up procedure.

The capacitor and interferometer are enclosed in an evacuat chamber in order to eliminate uncertainties in permittivity a refractive index. The permittivity, ϵ_0, in the formula is obtain from the permeability, μ_0, of vacuum and the velocity of light, μ_0 is by definition equal to $4\pi \times 10^{-7}$, and c is an expe mental constant now known[45, 46] to within about 3 parts in 1 With $c = 299\,792\,500$ m/s, ϵ_0 is $8{\cdot}854\,713 \times 10^{-12}$. When t figure is used in the formula, it is seen that the cross-capacitan is a little under 2 pF per metre length. The effective value the N.S.L. capacitor is about $\frac{1}{4}$ pF.

Anybody with experience of bridges will know that this val is uncomfortably small for resistance to be determined in ter of it and frequency to an accuracy within 1 part in 10^6, b here again the N.S.L. has gone a long way in achieving succe

In the first instance, capacitors of great constancy, with valu ranging from 1 pF to, say, $0{\cdot}1\,\mu$F in succession are measured terms of the calculable capacitance. This passage from small large capacitances is effected by transformer bridges, i.e. bridg in which the ratio arms are the two secondary windings of transformer, and the other two arms the known and unknow capacitors.[47] If one secondary arm has ten times as many tur as the other, the capacitances will be in the ratio of 1 : 10. very special precautions to ensure that the two secondary win ings enclose the same magnetic flux, the N.S.L. has succeed in making transformers with ratios accurate to 1 part in 1 Members of N.S.L. and N.B.S., working together at the N.B.S. have succeeded in making transformer bridges to meas capacitances ranging up to $1\,\mu$F. A frequency of 1 kc/s suitable, although other frequencies below 10 kc/s could be use

For the measurement of resistance in terms of capacitan and frequency the N.S.L. uses a parallel-arm bridge first describ by Ogawa. By switching to a Schering bridge, all the effects small phase defects in the components may be eliminated.

It is claimed[47] that by the methods described above, which re still being developed, especially as regards the production f stable capacitors and resistors of low temperature coefficient, should be possible to measure a resistance of 1 ohm to an ccuracy within 1 part in 10^6.

(3.4) Units of the Future

If the work described in the last two Sections comes up to xpectation, it will be possible in a few years to measure electric urrent in terms of the gyromagnetic ratio of the proton to an ccuracy within 1 part in 10^6 and electric resistance in terms of ength and time to the same accuracy. At present the gyro-agnetic ratio of the proton is a physical constant which must e measured in terms of current, length and time, and the first f these quantities cannot be determined from its definition to etter than 1 part in 10^5. If the devices developed for measuring ie gyromagnetic ratio in terms of current prove to be simpler, r at any rate not more troublesome, to use than a current alance, there might be something to say for abandoning the resent definition of the ampere and substituting for it a figure or the gyromagnetic ratio of the proton. The number chosen ay not, it is true, be 'correct' to better than 1 part in 10^5, but its doption would enable current to be measured with a repro-ucibility of 1 part in 10^6. Thus the two units required to build p the whole electromagnetic system would be obtainable to part in 10^6, whereas at present current at least is subject to an ncertainty ten times as great.

It is realized that a lot more work is needed before the General onference of Weights and Measures can be asked to examine ich a proposal, since even the much more convincing case of ie second of time has not yet been put before it, but the metre now all but defined in terms of a constant of nature, and it eems legitimate to bear in mind the likelihood of our other veryday units also being so defined in years to come.

(4) REFERENCES

(1) Hartshorn, L.: 'Units and Standards of Electrical Measurement. A Review of Progress', *Proceedings I.E.E.*, Paper No. 1403, November, 1952 (**99**, Part I, p. 271).

(2) Hartshorn, L.: 'Standards of Electrical Measurement. A Review of Progress', *Journal I.E.E.*, 1942 (**89**, Part I, p. 526).

(3) Essen, L., Parry, J. V. L., and Steele, J. McA.: 'Frequency Variations of Quartz Oscillators and the Earth's Rotation in Terms of the N.P.L. Caesium Standard', *Proceedings I.E.E.*, Paper No. 3002 M, August, 1959 (**107** B).

(4) Essen, L., Parry, J. V. L., Markowitz, W., and Hall, R. G.: 'Variation in the Speed of Rotation of the Earth since June 1955', *Nature*, 1958, **181**, p. 1054.

(5) Barrell, H., and Essen, L.: 'Atomic Standards of Length and Time', *Science Progress*, 1959, **47**, p. 209.

(6) *Procès-verbaux des séances du Comité International des Poids et Mesures*, 1956, **25**, p. 77.

(7) Essen, L., and Parry, J. V. L.: 'The Caesium Resonator as a Standard of Frequency and Time', *Philosophical Transactions of the Royal Society*, A, 1957, **250**, p. 45.

(8) Essen, L.: 'Atomic Clocks', *Journal I.E.E.*, 1958, **4**, p. 647.

(9) Markowitz, W., Hall, R. G., Essen, L., and Parry, J. V. L.: 'Frequency of Cesium in Terms of Ephemeris Time', *Physical Review Letters*, 1958, **1**, p. 105.

(10) Holloway, J., Mainberger, W., Reder, F. H., Winkler, G. M. R., Essen, L., and Parry, J. V. L.: 'Comparison and Evaluation of Cesium Atomic Beam Frequency Standards', *Proceedings of the Institute of Radio Engineers*, 1959, **47**, p. 1730.

(11) Kalra, S. N., Bailey, R., and Daams, H.: 'Canadian Caesium-Beam Standard of Frequency', *Nature*, 1959, **183**, p. 575.

(12) Essen, L., Parry, J. V. L., and Pierce, J. A.: 'Comparison of Caesium Resonators by Transatlantic Radio Transmission', *ibid.*, 1957, **180**, p. 526.

(13) Essen, L., Parry, J. V. L., Holloway, J. H., Mainberger, W. A., Reder, F. H., and Winkler, G. M. R.: 'Comparison of Caesium Frequency Standards of Different Construction', *ibid.*, 1958, **182**, p. 41.

(14) Kastler, A.: 'Les méthodes optiques d'orientation atomique et leurs applications', *Proceedings of the Physical Society*, A, 1954, **67**, p. 853.

(15) Bender, P. L., Beaty, E. C., and Chi, A. R.: 'Optical Detection of Narrow Rb^{87} Hyperfine Absorption Lines', *Physical Review Letters*, 1958, **1**, p. 311.

(16) *Procès-verbaux des séances du Comité International des Poids et Mesures*, 1958, **26**-B, p. M30.

(17) Essen, L.: 'Atomic Time and the Definition of the Second', *Nature*, 1956, **178**, p. 34.

(18) Essen, L.: 'Atomic and Astronomical Time', *ibid.*, 1956, **177**, p. 744.

(19) Essen, L.: 'Standards of Time and Frequency', *Research*, 1957, **10**, p. 217.

(20) Essen, L.: 'The Units of Time and Length', *Nature*, 1957, **180**, p. 137.

(21) National Physical Laboratory: 'MSF Standard Frequency Transmissions from the United Kingdom' (H.M. Stationery Office).

(22) National Physical Laboratory: 'Standard-Frequency Transmissions. Changes in the MSF Standard', *Electronic and Radio Engineer*, 1959, **36**, p. 117.

(23) Leclerc, G.: 'Rapport sur les comparaisons des étalons nationaux de résistance électrique effectuées en 1957', *Procès-verbaux des séances du Comité International des Poids et Mesures*, 1959, **26**–A, p. 101.

(24) Gautier, M.: 'Rapport sur les comparaisons des étalons nationaux de résistance électrique effectuées en 1955', *ibid.*, 1956, **25**, p. 120.

(25) Leclerc, G.: 'Rapport sur les comparaisons des étalons nationaux de force électromotrice effectuées en 1957', *ibid.*, 1959, **26**–A, p. 111.

(26) Gautier, M.: 'Rapport sur les comparaisons des étalons nationaux de force électromotice effectuées en 1955', *ibid.*, 1956, **25**, p. 128.

(27) Driscoll, R. L.: 'Measurement of Current with a Pellat-Type Electro-dynamometer', *Journal of Research of the National Bureau of Standards*, 1958, **60**, p. 287.

(28) Driscoll, R. L., and Cutkosky, R. D.: 'Measurement of Current with the National Bureau of Standards Current Balance', *ibid.*, 1958, **60**, p. 297.

(29) Driscoll, R. L.: 'Nuclear Magnetic Resonance and the Measurement of Magnetic Fields', National Physical Laboratory Symposium on Precision Electrical Measurements, Paper No. 8, November, 1954.

(30) Driscoll, R. L., and Bender, P. L.: 'Proton Gyromagnetic Ratio', *Physical Review Letters*, 1958, **1**, p. 413.

(31) Bender, P. L., and Driscoll, R. L.: 'A Free Precession Determination of the Proton Gyromagnetic Ratio', *Transactions of the Institute of Radio Engineers*, 1958, **I–7**, p. 176.

(32) Pake, G. E.: 'Fundamentals of Nuclear Magnetic Resonance Absorption', *American Journal of Physics*, 1950, **18**, p. 438.

(33) 'Research Highlights', Annual Report of the National Bureau of Standards, 1958, p. 3.

(34) Vigoureux, P.: 'An Absolute Determination of the Ampere', *National Physical Laboratory Collected Researches*, 1938, **24**, p. 173.

(35) Comité Consultatif d'Électricité: *Procès-verbaux des séances du Comité International des Poids et Mesures*, 1958, **26** B, p. E1.

(36) Rayner, G. H.: 'An Absolute Measurement of Resistance by Albert Campbell's Bridge Method', *Proceedings I.E.E.*, Monograph No. 95 M, April, 1954 (**101**, Part IV, p. 250).

(37) Rayner, G. H.: 'The Derivation of Resistance, Inductance and Capacitance from the N.P.L. Primary Standard of Mutual Inductance', *Transactions of the Institute of Radio Engineers*, 1958, **I–7**, p. 212.

(38) Henderson, J. T., and Romanowski, M.: 'A Standard of Mutual Inductance', *Canadian Journal of Physics*, 1955, **33**, p. 856.

(39) Romanowski, M., and Fraser, P. A.: 'Considerations on the Primary Winding of the Campbell Standard Inductor', *ibid.*, 1955, **33**, p. 871.

(40) Romanowski, M., and Olson, N.: 'An Absolute Determination of Resistance', *ibid.*, 1957, **35**, p. 1312.

(41) Thompson, A. M., and Lampard, D. G.: 'A New Theorem in Electrostatics and its Application to Calculable Standards of Capacitance', *Nature*, 1956, **177**, p. 888.

(42) Lampard, D. G.: 'A New Theorem in Electrostatics with Applications to Calculable Standards of Capacitance', *Proceedings I.E.E.*, Monograph No. 216 M, January, 1957 (**104** C, p. 271).

(43) Thompson, A. M.: 'The Cylindrical Cross-Capacitor as a Calculable Standard', *ibid.*, Paper No. 2887 M, May, 1959 (**106** B, p. 307).

(44) Annual Report of the National Standards Laboratory, Division of Electrotechnology, 1957–58, p. 5.
(45) ESSEN, L.: 'The Velocity of Propagation of Electromagnetic Waves derived from the Resonant Frequencies of a Cylindrical Cavity Resonator', *Proceedings of the Royal Society*, A, 1950, **204**, p. 260.
(46) FROOME, K. D.: 'Precision Determination of the Velocity of Electromagnetic Waves', *Nature*, 1958, **181**, p. 258.
(47) THOMPSON, A. M.: 'The Precise Measurement of Small Capacitances', *Transactions of the Institute of Radio Engineers*, 1958, I–7, p. 245.
(48) MCGREGOR, M. C., HERSH, J. F., CUTKOSKY, R. D., HARRIS, F. K., and KOTTER, F. R.: 'New Apparatus at the National Bureau of Standards for Absolute Capacitance Measurement', *ibid.*, 1958, I–7, p. 254.
(49) OATLEY, C. W., and YATES, J. G.: 'Bridges with Coupled Inductive Ratio Arms as Precision Instruments for the Comparison of Laboratory Standards of Resistance or Capacitance', *Proceedings I.E.E.*, Paper No. 1631 M, March, 1954 (**101**, Part III, p. 91).
(50) CLOTHIER, W. K.: 'A Fixed Gas-Dielectric Capacitor of High Stability', *ibid.*, 1954, **101**, Part II, p. 453.
(51) ROMANOWSKI, M., and BAILEY, R.: 'Suggested Modifications to a Method for the Determination of the Absolute Ampere', *Canadian Journal of Physics*, 1959, **37**, p. 896.
(52) DUNN, A. F.: 'Determination of the Unit of Capacitance', *ibid.*, 1959, **37**, p. 35.
(53) RAYNER, G. H.: 'The Calibration of Inductors at Power and Audio Frequencies', *Proceedings I.E.E.*, Monograph No. 315 M, October, 1958 (**106** C, p. 38).
(54) STEELE, J. MCA.: 'The Standard Frequency Monitor at the National Physical Laboratory', *ibid.*, Paper No. 1765 M, October, 1954 (**102** B, p. 155).
(55) LAW, H. B.: 'Standard Frequency Transmission Equipment at Rugby Radio Station', *ibid.*, Paper No. 1762 R, October, 1954 (**102** B, p. 166).
(56) ARDITI, M., and CARVER, T. R.: 'Optical Detection of Zero-Field Hyperfine Splitting of Na', *Physical Review*, 1958, **109**, p. 1012.
(57) HARTSHORN, L.: 'Precision Electrical Measurements', *Research*, 1959, **12**, p. 307.

SURFACE WAVES: A PROPOSED DEFINITION

By Prof. H. E. M. BARLOW, Ph.D., B.Sc.(Eng.), Member.

(*Communication received 2nd November*, 1959.)

In approaching this problem it is important to bear in mind the considerations that have led in the past to the definitions now universally accepted of other forms of guided electromagnetic wave. All these definitions are based on behaviour in idealistic conditions. Thus the so-called TEM wave associated with twin parallel conductors assumes a lossless guide, and consequently such a wave cannot exist in practice. The waveguide modes inside smooth-bore hollow metal tubes are identified in relation to perfectly conducting walls and with the sole exception of the circular H_{0n} and E_{0n} families can have no separate existence in an actual waveguide. Having defined these wave modes as recognizable field configurations supported by ideal media, we can then go on to discuss practical cases involving a variety of perturbations, sometimes of quite drastic character. There is no difficulty in doing this and no confusion arises from it. In fact the procedure leads to a much simpler and better approach to a common understanding of guided wave modes than could possibly be achieved starting from wider definitions in the first place.

The really characteristic feature of the surface wave is its non-radiating property, and this should therefore form the basis of its definition. Accepting that feature as the essential one, it follows that the supporting surface must be straight in the direction of propagation of the wave so that when curved surfaces occur they must be regarded as perturbations. Just as we speak of a TEM wave supported by a parallel-strip line, so we can discuss with equal justification the azimuthal surface wave on a cylindrical surface, and this will be interpreted without ambiguity as having a close approach to the evanescent field structure of the pure surface wave above a flat surface. We can also talk of surface-wave aerials on the same kind of basis, recognizing that such an arrangement might consist, for example, of a single-wire transmission line with discontinuities deliberately introduced at intervals along its length. Any attempt to include a variety of perturbed field distributions within the basic definition of the 'surface wave' is clearly fraught with great difficulty and can only lead to further confusion.

It is suggested, therefore, that we define a surface wave in the following very simple terms:

> A *surface wave* is one that propagates along an interface between two different media without radiation, such radiation being construed to mean energy converted from the surface-wave field to some other form.

A physical definition of this kind would, it is felt, do much at the present time to avoid misunderstandings between those concerned with surface-wave problems, and moreover such a definition is susceptible of precise mathematical interpretation.

Since the definition is an idealistic one, realizable in practice only with a perfectly smooth and longitudinally straight interface between homogeneous media, no qualifying statements are necessary, as, for example, that the wave should be associated with finite energy. The homogeneous plane wave is defined without any such restriction.

The fact that an exact mathematical solution exists for what we have called a perturbed surface wave propagating in azimuth around a cylindrical supporting surface does not detract in any way from the argument in favour of the proposed definition based on a complete absence of radiation. The situation is similar to that of the circular H_{0n} and E_{0n} families of waveguide modes in a hollow tubular guide with a finite wall impedance. The field distribution within the guide is modified over the cross-section by the surface impedance, but it remains basically of the same pattern.

We conclude, therefore, that the best procedure in defining a surface wave is to extend the traditional approach already applied to other wave modes now well recognized, and in pursuance of that view to regard the pure surface wave simply as one that propagates without radiation along an interface between two different media.

Prof. Barlow is Pender Professor of Electrical Engineering, University College, London.

1.396.2: 621.396.65

The Institution of Electrical Engineers
Paper No. 3033 E
Nov. 1959

FREQUENCY PATTERNS FOR MULTIPLE-RADIO-CHANNEL ROUTES

By B. B. JACOBSEN, B.Sc.(Eng.), Member.

he paper was first received 10*th August*, 1957, *and in revised form* 3*rd April*, 1959. *It was published in November*, 1959, *and was read before the* ELECTRONICS AND COMMUNICATIONS SECTION 7*th December*, 1959.)

SUMMARY

In radio trunk systems a number of two-way radio channels are ed over a common route with repeaters provided at intervals. Each dio channel carries a baseband signal by frequency modulation of e radio carrier frequency. The baseband signal may consist of an sembly of single-sideband telephony channels and/or a television annel.

When choosing the frequency pattern for the radio channels, lectivity requirements are found to arise, not only from the primary ct that the channels are adjacent in the frequency spectrum, but also om secondary features of the equipment such as, for instance, the termediate frequency used in the heterodyne repeaters.

The paper considers the effects which give rise to such selectivity quirements and shows how to choose the channel frequency pattern d the intermediate frequency to minimize the selectivity require-ents. The choice takes account of amplitude non-linearity in wave-ide circuits. The presence of this was expected on theoretical grounds d has been confirmed by measurements.

A specific frequency pattern is worked out for a route with six o-way channels in which each radio channel can carry at least 600 ephone channels or a single television channel. This arrangement as adopted by the C.C.I.R. in 1956.

LIST OF SYMBOLS

S = Channel frequency spacing, Mc/s.
G = Extra frequency spacing between send and receive groups, Mc/s.
IF = Intermediate frequency.
F = Shift frequency, the frequency difference between the signal input and output frequencies of a repeater.
LO_1 = Local-oscillator frequency used at the input mixer of the repeater.
LO_2 = Local-oscillator frequency used at the output mixer of the repeater.
Brr = (r.m.s. radian)^2 phase deviation in decibels with respect to (1 rad r.m.s.)^2.
B = Maximum useful baseband frequency.

(1) INTRODUCTION

Broad-band radio-relay systems are being used to supplement ble transmission systems in the long-distance telephone net-ork.[1]

A single radio channel can handle a large number of single-leband telephone channels and/or a television channel. Two dio channels in opposite directions but along the same route ovide the two-way transmission system.

The input (or baseband) signal is carried, on the radio channel, frequency modulation of the radio carrier frequency at the ansmitting terminal. Radio transmission is by line-of-sight ths of approximately 30 miles in length and the radio fre-encies used are higher than 1 000 Mc/s. At the end of a radio ansmission section, the radio signal is amplified in a repeater d retransmitted at a somewhat different radio frequency over the next section; generally, a particular radio channel uses only two frequencies. After the required number of sections, the baseband signal is recovered from the radio signal at the receiving terminal. Several radio channels are usually needed over a common route, and such a group of channels share much of the capital cost of the route, particularly station sites, buildings and approach roads and also aerials, feeders, maintenance equipment and power supplies.

A very important advantage of having a group of radio channels is that a spare channel becomes an economical possibility. A standby channel, even if shared, gives a considerable measure of protection, not only against equipment failure but also against selective fading in a working channel.[2]

Radio routes are likely to be used across international boundaries, and for this reason it is desirable that certain characteristics of the radio channels shall be internationally agreed. The Comité Consultatif International Radio (C.C.I.R.) is the international body responsible for this field, and this committee has studied radio relay systems of the type under discussion; its recommendation No. 194, 'Standardization of Multichannel Radio Relay Systems using Frequency Division Multiplex, has been published[3] and covers the pattern of frequencies for use on a radio route providing six radio channels in each direction of transmission in a total bandwidth of about 400 Mc/s. Each radio channel is intended to carry 600 telephone channels, but the same pattern of frequencies is also recommended when radio channels are required for transmission of television signals (C.C.I.R. Recommendation No. 195).

The present paper gives in detail the considerations that led to the choice of the recommended frequency pattern.

An earlier C.C.I.R. proposal for such a frequency pattern revealed certain difficulties at the design stage. It was then considered desirable to study the problem of frequency patterns of multi-radio channel systems sharing a route to see if a more satisfactory pattern could be found and if at the same time certain desirable new features could be introduced. Section 2 gives the terms of reference of this study, and later Sections show how the preferred frequency pattern was obtained.

(2) INITIAL SPECIFICATION FOR THE NEW FREQUENCY PATTERN

In the systems in Reference 1, the frequencies used in a section for the two directions of transmission were interleaved, but in the earlier C.C.I.R. proposal, already referred to, the frequencies for the two directions of transmission were placed in separate groups. The latter arrangement is preferred because it simplifies the transmitter-to-receiver coupling problems and permits an important saving in the bandwidth required. The general plan for such a frequency arrangement is shown in Fig. 1.

Two repeater stations, r and $(r + 1)$, are shown. Station r receives the higher-frequency group of channels from either direction and converts them, before retransmission, to the lower-frequency group, as indicated diagrammatically by the arrows. In station $(r + 1)$, and also in station $(r - 1)$ which is not shown, an upward frequency translation is used, as indi-

Mr. Jacobsen is with Standard Telephones and Cables, Ltd.

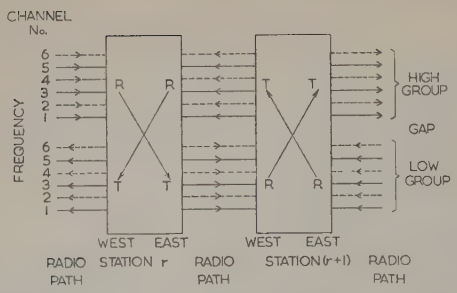


Fig. 1.—Multi-channel radio route, showing two repeater stations. Grouped transmission.

cated by the arrows. Thus, along a route the repeaters alternate between the two general types illustrated in Fig. 1.

The requirements detailed below, which were fixed before beginning the study, were intended primarily to apply for frequency-modulated r.f. channels, but they may equally well be suitable for amplitude-modulated channels. They are:

(i) Six two-way radio channels are required, arranged as shown in Fig. 1.

(ii) A bandwidth of about 400 Mc/s can be made available.

(iii) The intermediate frequency at the receiver and repeaters is to be 70 Mc/s unless there are strong reasons for a different value.

(iv) Each radio channel is to be suitable to carry at least 600 telephone channels (or a television channel).

(v) It is very desirable that two sets of frequency allocations, (*a*) and (*b*), shall be made available.

In Fig. 1, the frequencies on both sides of the repeaters are shown as equal. It will in some cases be desirable for the frequencies on the east side to be chosen midway between the west-side frequencies (interleaved or mutually staggered).

This requirement arises when one set of frequencies intended, for instance, for the east side of the repeater reaches the west side by reflection from trees or buildings which are in the foreground of the west-side aerial and which are illuminated by the remote transmitter associated with the east side of the repeater.

(vi) The frequency pattern shall as far as possible be so chosen that the selectivity requirements are controlled by the basic effects (such as the presence of adjacent channels) rather than by secondary effects such as, for instance, image responses.

(vii) It is desirable that an initial installation of up to three two-way channels shall require only a single aerial and feeder for each direction. This entails using one feeder and aerial for both transmission and reception. The remaining three two-way channels similarly shall depend on a second aerial. The purpose of this part of the specification is to improve the economy of small (initial) installations, and to reduce the incidence of disturbance caused when installing further channels.

(viii) There are considerable advantages in using polarizations differing by 90° for adjacent channels. Cross-polarization provides selectivity without filtering. Relative cross-polarization (CP) between adjacent channels should therefore be specified.

(3) REPEATER EQUIPMENT

The type of repeater which will be considered in working out the frequency pattern is shown in Fig. 2, which gives the equipment required for the three odd-numbered two-way channels [Section 2, requirement (vii)]. A similar amount of equipment would be required for the even-numbered channels. Odd- and even-numbered channels would use relative cross-polarization [Section 2, requirement (viii)].

In Fig. 2, each aerial feeder terminates in a filter group which provides for three channels in each direction of transmission. The equipment required for a one-way repeater is shown in detail only for Channel 1 in the west-to-east direction. It comprises the low level mixer, LLM, in which the incoming signal is modulated by the frequency LO_1 whereby the signal is translated to the intermediate-frequency range for amplification in

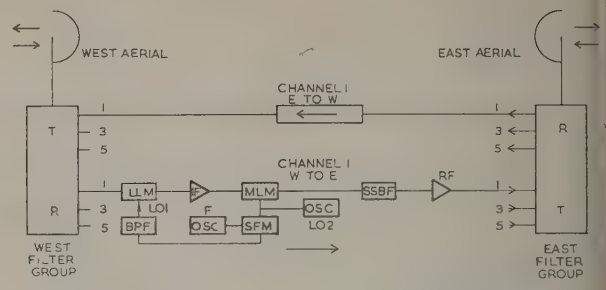


Fig. 2.—Repeater for three two-way radio channels.

SFM Shift frequency mixer.
LLM Low-level mixer.
IF Intermediate-frequency amplifier.
MLM Medium-level mixer.
RF Radio-frequency amplifier.
SSBF Single-sideband filter.

the i.f. amplifier. Following this, the i.f. signal is modulated by the frequency LO_2 in the medium-level mixer, MLM, in which the signal is translated to the required output range. A filter, SSBF, is provided to attenuate unwanted products from MLM but to pass the required single-sideband output signal to the radio-frequency amplifier, RF, and then to the east filter-group and to the east aerial.

The local-oscillator frequency, LO_1, is produced from LO_2 by modulation by the 'shift frequency', F, which will be of the order of 200 Mc/s. The exact value of the shift frequency depends on the relationship required between the frequencies on the two sides of the repeater.

The general outline of the plan to be specified is shown in Fig. 3. The quantity S is the channel carrier-frequency spacing.

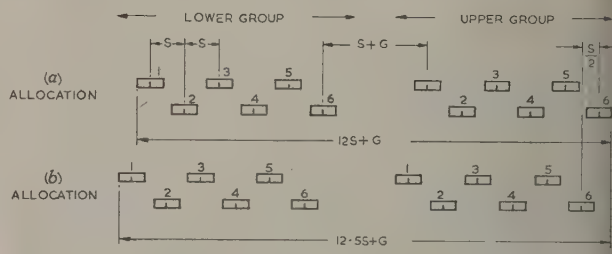


Fig. 3.—Frequency plan.

G is the extra spacing between the send and receive channel frequencies. Two allocations, (*a*) and (*b*), are shown, (*b*) being $S/2$ below the (*a*) allocation. Either of the allocations might be used in a given transmission section, i.e. might apply to one side of a repeater. Normally, channel frequencies from the two allocations will not be mixed in the same transmission section.

The total bandwidth required for the two allocations is $12{\cdot}5S + G$. This assumes that the boundary channels are of width $\pm S/2$, which is somewhat greater than the real bandwidth.

The shift frequency required will depend on the frequency allocations used on the two sides of the repeater. If the same allocation, (*a*) or (*b*), is used the shift frequency is $F = 6S + G$. If the (*a*) allocation is used on one side and (*b*) on the other, the shift frequencies will be different for the two directions of transmission and the values will be $F + S/2$ and $F - S/2$ respectively.

In choosing local-oscillator frequencies for a one-way repeater for a channel, there are four possibilities. LO_1 may be either above or below the incoming frequency, and LO_2 may similarly

e above or below the outgoing frequency. Only two of these rrangements are preferred, namely those in which the two local-scillator frequencies are both either above or below their sociated frequencies. These preferred arrangements have the dvantage that the input and output frequencies differ exactly y the frequency F (see Fig. 2) independently of the actual alue of LO_2, which need therefore not be very stable. Error n the repeater channel-frequency-shift will then arise only from ne frequency F rather than from errors in the much higher requency, LO_2. These arrangements may be said to give stable hift.

If the frequency LO_1 is above the incoming frequency, the esulting intermediate frequency will have inverted sense of requency modulation as compared with the input frequency, ut if LO_1 is below the incoming frequency, there will not be nversion. The output frequency in both cases will have the ame sense of modulation as the input frequency when the stable shift' arrangement is used.

There are two advantages in having inverted sense of i.f. nodulation in alternate repeaters.

First, if the i.f. equipment has a systematic, as opposed to a andom, component of group-delay error which is linear with requency, this will be cancelled out between a pair of repeaters sing opposite senses of frequency deviation in the i.f. equipment.

Secondly, if the i.f. equipment has a systematic component of roup-delay error which is proportional to the square of the lifference between the actual i.f. signal-frequency and the equip-nent centre-frequency, then an error, Δ, in the signal mean-requency will cause the signal to suffer a group-delay error which is linear with frequency and proportional to Δ. This effect will arise if the mean frequency from the terminal equip-nent is in error, but if the i.f. signals in alternate repeaters are nverted, the effects will be cancelled out, since alternate Δ values will be positive and negative.

This arrangement (balanced i.f. inversion) gives an improve-nent only when the i.f. equipment has systematic group-delay errors. Systematic errors may be present even when the group delay has been approximately equalized, but the residual errors nay then be so small that the balanced i.f. inversion advantage s not required.

Balanced i.f. inversion could therefore be an aim, but not an essential one. To attain it, successive repeaters should be arranged with both local-oscillator frequencies alternately above and below the associated transmission frequencies.

In Fig. 1 it may be seen that alternate repeaters are of the type which increases (as in $r + 1$) or which decreases (as in r) the channel frequencies when they pass through the repeater. This requirement, and the aim of having balanced i.f. inversion, reduces the choice of local oscillator arrangements to that of coupling the two sets of alternatives. For repeater r, for instance, both the local-oscillator frequencies for a one-way channel could be chosen to be below the associated transmission frequencies; for the following repeater, both local-oscillator frequencies must then be above. Fig. 4 shows this arrangement for

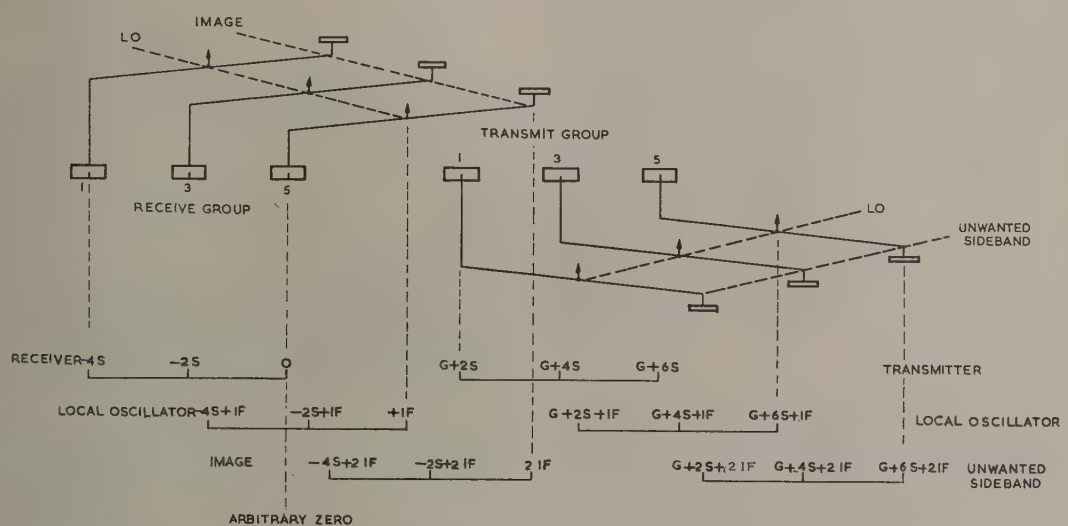


Fig. 4.—Location of local-oscillator frequencies, image responses and unwanted sideband. Preferred arrangement.

repeater $r + 1$, and Fig. 5 shows the other possible choice. In these diagrams, the frequency scale is horizontal but the frequency bands associated with the incoming and outgoing signals, the local-oscillator frequency, and the image and unwanted sideband frequencies have been separated vertically to give a clearer presentation.

In the arrangement in Fig. 4, it will be seen that the image response of the receiver falls towards or into the transmitting group used for the other direction of transmission. It might be thought that this would lead to high-selectivity requirements for the receiver, but a receiving filter is necessary in any case to control a number of effects which require discrimination (e.g. the adjacent channels). By choosing a favourable pattern in which the receiver image-frequencies fall between the transmitter frequencies (which use the same aerial), the receiving-filter requirements will not be increased by the present effect.

In the arrangement in Fig. 5, the unwanted sideband which arises in the transmitting equipment falls towards or into the receiving group of the other direction of transmission, but the image response falls outside the frequency allocation for the route under consideration.

When transmitters and receivers share an aerial, the unwanted sideband in Fig. 5 will require more suppression if it falls in or near a receiving-channel frequency than if it falls in the frequency range of other transmitting channels or outside the frequency band of the group of systems using the route. This is a disadvantage.

In Fig. 5, the image-frequency region is exposed to signals

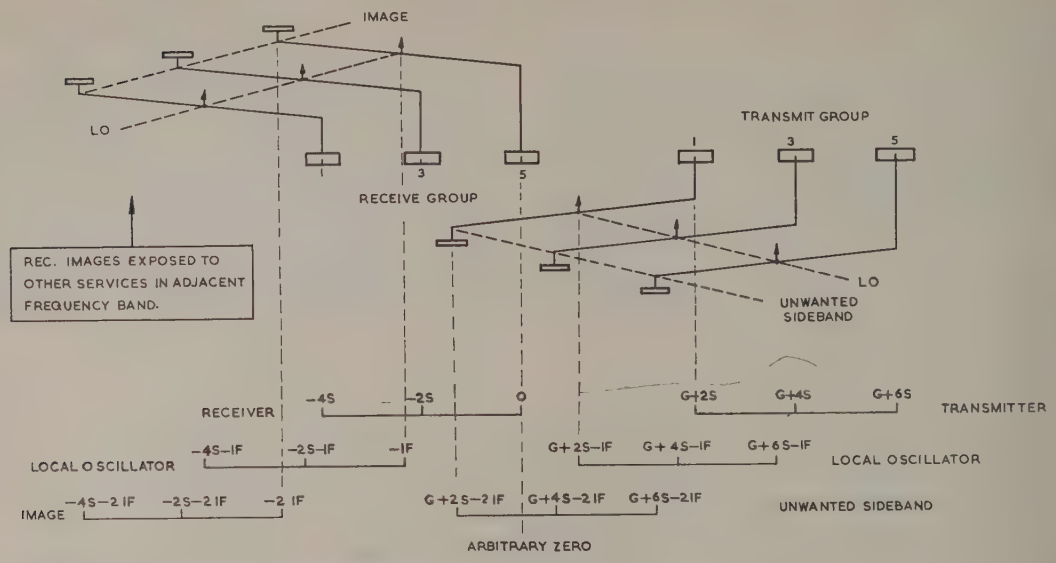


Fig. 5.—Location of local-oscillator frequencies, image responses and unwanted sideband. Alternative arrangement not used.

belonging to other services using frequencies which are not controllable or even known and which might use a high power. This is another disadvantage.

The arrangement of Fig. 4 is therefore preferred. The preferred arrangement is characterized by the choice, for the receive local oscillation LO_1, of frequencies which lie towards the transmitting group of frequencies. At a station of type *r* (Fig. 1) the frequency pattern will be that of Fig. 4 as seen in a mirror. The patterns for the even-numbered channels are similar to those shown, but shifted by an amount *S*.

The general pattern of frequencies used in a repeater is now clear in principle. Before proceeding to the main task of determining the optimum values for the frequency parameters *S*, *G* and *IF* which ensure minimum filtering requirements, it is necessary to consider:

(*a*) The frequency regions in which the receivers, without filters, are most sensitive to interference (Section 4).
(*b*) The sources of unwanted signals which are present in the equipment shown in Fig. 2 and their frequencies and levels, without filters (Section 5).
(*c*) The frequency differences between each sensitive region of each receiver and each source of unwanted signal (individual couplings). These differences (or clearances) will be expressed in terms of *S*, *G* and *IF*, and the final work in the further Sections will be to find numerical values for these parameters which maximize simultaneously all clearances for couplings which could cause controlling filtering requirements. The couplings are mainly those which occur via the filter groups in Fig. 2, i.e. between the three transmitters working in one direction and receivers working in the opposite direction of transmission, using a common aerial.

(4) SENSITIVE REGIONS OF THE RECEIVER

In this Section it will be assumed that the equipment shown in Fig. 2 is used to convey a carrier frequency which is phase or frequency modulated by a baseband signal. At the end of a chain of such stations, the baseband signal is recovered from the modulated carrier. The effect of harmful spurious input frequencies to repeaters will manifest itself as noise in the recovered baseband signal, and the noise will be characterized by its magnitude and frequency.

At a particular repeater, a spurious signal which enters the input circuit will pass through the mixer and the i.f. amplifier. The i.f. equipment will generally contain a limiter or its equivalent and in this the spurious signal will be incorporated into the f.m. wave in the form of 'phase noise' (see Reference 4, Section 18.1).

Phase noise is a spurious phase modulation of the required carrier wave. It occurs at a baseband frequency equal to the difference between the frequencies of the carrier and the spurious input. Spurious frequencies at later repeaters are similarly incorporated in the carrier signal. When a spurious input occurs at a frequency which differs from the carrier frequency by an amount somewhat greater than the maximum baseband frequency, the resulting phase noise will not disturb the circuit so much, since it largely falls outside the useful baseband range. There is thus a 'sensitive range' in the neighbourhood of the wanted carrier frequency.

If the spurious signal differs sufficiently from the wanted carrier, it will be attenuated after the mixer by the i.f. equipment and be correspondingly less dangerous. In discussing these sensitive ranges, it is preferable to discuss the sensitivity in terms of power at the mixer, i.e. independently of any filter effects, since these have not, in fact, been fixed. The i.f. selectivity may, however, with advantage be included.

There are other sensitive ranges of which the image response is the best known. The image-response range is exactly similar to the main response; it occurs at a frequency distant $2IF$ from the wanted carrier and distant IF from the local oscillator frequency, LO_1, of the receiver, as has already been indicated in Fig. 4.

The local-oscillator frequency region also has a response to spurious input; this arises because the effective local oscillation will be phase modulated in the presence of the spurious signal, and consequently the i.f. signal will also be phase-modulated. Owing to the high level of the local oscillation LO_1 (usually about 1 mW), the sensitivity of the local-oscillation region will be smaller than that of the image and main regions. If the mixer is balanced, the sensitivity to interference will be reduced, depending on the degree of balance attained.

These responses are shown in Fig. 6. The vertical scale indicates sensitivity to disturbance, while the horizontal scale

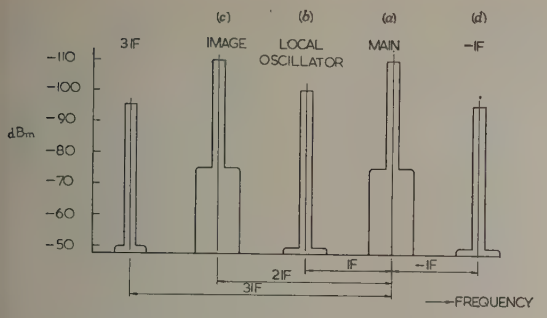


Fig. 6.—Sensitive regions of a repeater or receiver.

licates frequencies relative to the wanted carrier and local-cillation frequencies. Some further responses are also shown. ose marked *IF* and 3*IF* have the same mechanism of genera-n. Spurious signals in either of these ranges will phase odulate the effective local oscillation by a frequency $2IF + \Delta$; e intermediate-frequency signal will therefore also be phase odulated by $2IF + \Delta$, before the i.f. selectivity has been fective. This phase modulation results in two sidebands of nich one, $IF - 2IF - \Delta$, falls back into the i.f. range and is nplified there in the same way as a spurious signal in the main nge, and is incorporated as phase noise at frequency Δ.

The paper is not concerned with the calculation of the per-ssible amount of spurious signal, but rather with frequency tterns which do not expose the most sensitive ranges to urious signals.

In Fig. 6 the sensitivity figures are given as an example and e expressed by simple straight lines; in fact, the actual sen-ivity near the middle of each range is, in most systems, much naller than is shown.

It may at first sight seem surprising that the levels of tolerable terference in Fig. 6 are expressed in absolute values instead being related to the received wanted signal. When, however, is remembered that one of the principal sources of noise in a nk is thermal agitation, it is easily seen that it is both proper d convenient to relate the acceptable amount of an interfering equency to the basic thermal-agitation noise power which is perative at the input to a receiver.

The effect of interfering frequencies may therefore be expressed an equivalent increase in the noise factor of the receiver. nis method is particularly appropriate when the disturbed rrier frequency is 'artificially dispersed'. Section 13 shows by ay of example how the main-region sensitivity in Fig. 6 may estimated.

The width of the most sensitive region is slightly more than *B*, where *B* is the maximum baseband frequency of the trans-ission system, i.e. of the order of $\pm 2{\cdot}5$ Mc/s for a 600-channel l.m. carrier telephone system or ± 5 Mc/s for a 625-line tele-sion system. An extra allowance, needed when the disturbed, id/or disturbing carrier signal is itself frequency-modulated by gnals, has not been made in Fig. 6, which therefore indicates e 'net clearance' required of spurious signals in order to ensure at the severe noise effects do not arise in the received base-nd frequency range. In Fig. 6, the sensitive regions are indi-ted to have a greater width at a lower level of sensitivity. nis corresponds to the fact that there are other mechanisms of oducing baseband noise which will not be further considered the paper. The bandwidth of these less sensitive ranges rresponds to the bandwidths of the i.f. equipment used.

To summarize this Section very briefly: spurious signals which ay reach the input mixer of a repeater (or receiver) should preferably not be located with less than $\pm B$ Mc/s 'net clearance' of the following frequencies, relative to the wanted carrier:

	Designation
Channel carrier frequency	(*a*)
Local-oscillation frequency, $+IF$ ..	(*b*)
Image region, $+2IF$	(*c*)
$-IF$ region	(*d*)
$+3IF$ region	—

(5) SOURCES OF DISTURBING SIGNALS

All high-power levels in the repeater of Fig. 2 are potentially harmful and must be attenuated in all paths leading to an input mixer. The most obvious high-level source is a transmitting channel, but equally important is the unwanted sideband from the mixer *MLM* (Fig. 2). These are both at full transmitting level, but some suppression of the unwanted sideband must be provided by the filter SSBF, irrespective of overall considerations. The frequency LO_2 will also be present at the output of the amplifier RF. In spite of balancing in the mixer MLM, the level may be approximately equal to full signal level (apart from the effects of filter SSBF). Each transmitting channel provides these three terms.

Another source of potentially disturbing frequencies is the receiver LO_1. From one receiver, the frequency LO_1 may travel via the filter group to another receiver, or to the aerial. There are three such frequencies at each end of the equipment shown in Fig. 2.

(5.1) A Special Source of Disturbing Frequencies

When two or more transmitters are using a common waveguide circuit, it is theoretically possible that combination frequencies are produced by amplitude non-linearity in the waveguide circuits. Tests on actual equipment confirmed the presence of such effects. Two unmodulated transmitting frequencies, *A* and *B*, of power 5 watts each were found to produce the frequency $2A–B$. The level of the product term was found to vary considerably between various waveguide circuit components which were tested. It was found to vary with the surface condition at joints in the waveguide. After many measurements it was decided that it would be safe to assume that, in practice, the combination terms without excessive precautions, would not, at 4 Gc/s, exceed a level of -80 dBm for 5 watt transmitting power. A disturbing level of this magnitude could not be tolerated in the main response range of a receiver (Fig. 6).

The non-linear effects will arise in parts of the waveguide circuit in Fig. 2 which are common to both transmit and receive channels, and if the product term occurs near a receiving frequency it is not possible to reduce the interference by filtering. The only way of making the amplitude non-linearity type of interference acceptable is to select a frequency pattern in which the interference terms fall clear of the main response range (Fig. 6).

All other interference terms can be controlled by providing selectivity, and in their case the object of choosing a frequency pattern is only to minimize the selectivity requirements. The non-linear effect, therefore, has a first claim on benefits obtainable from the choice of the frequency pattern.

It may be noted that certain distortion terms arise in the modulation process in the shift-frequency mixer, SFM, in Fig. 2, which, in unfavourable cases, cannot be filtered out but will reach LLM; their levels can, however, be adequately controlled by design of the mixer SFM. One such unfavourable case occurs when a harmonic of the shift frequency, *F*, falls near the required frequency, LO_1.

The following will be used to designate the sources of disturbance:

Main transmit channel	A
Transmit local oscillation	B
Unwanted sideband	C
Receive local oscillation	D
Third-order non-linear products ..	E

(6) PRINCIPAL UNWANTED COUPLINGS

Section 4 discussed the sensitive regions of a receiver and Section 5 listed the output frequencies from a transmitter. Each two-way channel has two such sets. In Section 3 a particular repeater arrangement was preferred, in which LO_1 will be above the receive channel when this is in the low group, but below it when the channel is in the high group; the transmitting local-oscillator frequency LO_2 follows the opposite rule.

The chosen arrangement is shown in Fig. 4 for one side of the repeater ($r + 1$), and the diagram applies equally for either the odd- or the even-numbered channels and for both the (*a*) and (*b*) frequency allocations. (No absolute frequency is defined in Fig. 4.)

If in Fig. 4 the frequency scale is inverted, the resulting pattern is that required for repeater r in Fig. 1. It is, therefore, only necessary to formulate the clearances in terms of S, G and IF with reference to Fig. 4. This has been done in Table 1.

The first two lines of the Table give the sources of disturbances A to E (Section 5), and the power which would be transmitted to the filter group if no filters were present (this is purely for guidance). In the next line are shown the frequencies at which each type of source occurs; these are read from Fig. with respect to the arbitrary reference point shown.

The first column shows the designations of the various sensitiv regions (Section 4) with the addition of (*e*), which refers to th channel receiving frequency ranges of the adjacent repeat station (which, of course, agree in frequency each with one the A-frequencies). The next column gives, purely for guidanc the power which might be tolerated in each of the sensitiv ranges, assuming that there is an exposure at each repeater poi in a long circuit. The third column gives the frequencies of th sensitive region, read from Fig. 4.

In the body of the Table are shown the 'gross clearances' couplings which arise. A gross clearance is the differenc between the centres of individual disturbing and disturbe ranges. In the Table, only the extreme values of each type coupling are shown, but intermediate values also occur, o for each integral multiple of $2S$ between the extremes shown.

To make the derivation of the Table clearer, consider the fir square: this refers to couplings between sources A and sensitiv regions *a*. There are nine such couplings A*a*, three frequenci A and three ranges *a*). The largest clearance is between A frequency $G + 6S$ and *a* at frequency $-4S$ and this give $G + 10S$. The smallest clearance is between A at frequenc

Table 1

GROSS CLEARANCES

(The clearance is the numerical value of the quantity shown in the body of the Table)

Region ⟶ Disturbing / ↓ Disturbed	Power dBm → Output ↓ Tolerable	Relative frequency → ↓	A (Transmit channel)	B [LO_2]	C (Unwanted sideband)	D [LO_1]	E (Non-linear effects in W.G.)
			+37	+37	+37	0	−80
			$G \begin{cases} +2S \\ +4S \\ +6S \end{cases}$	$G + IF \begin{cases} +2S \\ +4S \\ +6S \end{cases}$	$G + 2IF \begin{cases} +2S \\ +4S \\ +6S \end{cases}$	$+IF \begin{cases} -0 \\ -2S \\ -4S \end{cases}$	$G \begin{cases} -2S \\ +0 \\ +2S \end{cases}$
a (Receive channel)	−110	0 $-2S$ $-4S$	O $G + 2S$ to $G + 10S$	O $G + IF + 2S$ to $G + IF + 10S$	O $G + 2IF + 2S$ to $G + 2IF + 10S$	I $IF - 4S$ to $IF + 4S$	IV $G - 2S$ to $G + 6S$
b (Near LO_1)	−100	$IF \begin{cases} -0 \\ -2S \\ -4S \end{cases}$	II $G - IF + 2S$ to $G - IF + 10S$	O $G + 2S$ to $G + 10S$	O $G + IF + 2S$ to $G + IF + 10S$	—	M $G - IF - 2S$ to $G - IF + 6S$
c (Near image)	−110	$2IF \begin{cases} -0 \\ -2S \\ -4S \end{cases}$	III $G - 2IF + 2S$ to $G - 2IF + 10S$	II(*a*) $G - IF + 2S$ to $G - IF + 10S$	O $G + 2S$ to $G + 10S$	M $-IF - 4S$ to $-IF + 4S$	—
d (Near $-IF$)	−95	$-IF \begin{cases} -0 \\ -2S \\ -4S \end{cases}$	O $G + IF + 2S$ to $G + IF + 10S$	O $G + 2IF + 2S$ to $G + 2IF + 10S$	O $G + 3IF + 2S$ to $G + 3IF + 10S$	V(*a*) $2IF - 4S$ to $2IF + 4S$	M $G + IF - 2S$ to $G + IF + 6S$
e (Receive channel at next station)	−110*	$G \begin{cases} +2S \\ +4S \\ +6S \end{cases}$	$-4S \quad -2S \quad 0$ $-2S \quad 0 \quad +2S$ $0 \quad +2S \quad +4S$	I(*a*) $IF - 4S$ to $IF + 4S$	V $2IF - 4S$ to $2IF + 4S$	II(*b*) $IF - G - 10S$ to $IF - G - 2S$	—

* The radio transmission-path loss has not been allowed for.

The first two lines of the Table give the sources of disturbances A to E (Section 5), and the power which would be transmitted to the filter group if no filters were present (this is purely for guidance). In the next line are shown the frequencies at $G + 2S$ and *a* at frequency 0, which gives $G + 2S$. This parti cular type of coupling gives good clearances, which do no reach low values. The clearance for A*a* is not critically dependen on the precise values of G, S and IF, and will always be large The letter O in a square indicates adequate spacing.

Consider now coupling between sources A and sensitive range

The clearances in this case range from that between the frequencies $G + 6S$ and $-(4S - 2IF)$, which is $G - 2IF + 10S$, and that between $G + 2S$ and $2IF$, or $G - 2IF + 2S$; the latter clearance will generally be negative, and since the sign of the clearance is for most purposes of no importance, it is the magnitude which should be considered, as indicated in the title of the Table. In the particular term Ac which is now under consideration, the clearance is therefore $G - 2IF$ $n2S$ with n having values 5, 4, 3, 2, 1. There are nine terms but only five different values of clearance. The minimum clearance value is the only one of importance for the present purpose, and this value for Ac type of coupling is critically dependent on the choice of G, S and IF. The particular square under discussion is marked III, and if the clearance for type Ac coupling is not bigger than the maximum baseband frequency, the amount of filter discrimination required will, according to the Table, be 147 dB. If, however, the net clearance is greater than the maximum baseband frequency, a discrimination of perhaps 110 dB will be sufficient. If the clearance is greater than the i.f. equipment bandwidth, the discrimination required will be even less. The clearance values calculated above are 'gross clearances'; it is necessary in practice to allow for the effect of frequency modulation, and this will be discussed later in connection with 'net clearances'.

By considering all the squares in Table 1 it is found that there are several in which small clearances could occur which would demand a large filter discrimination. These have been marked with roman numerals. Some squares are marked to denote minor effects for which adequate filter discrimination will be available. In one square, Ae, all the clearance figures have been filled in. The line of zero clearances are the only wanted couplings in the Table as they represent the intended transmission; the remaining figures in Ae are clearances for coupling to the remote repeater.

For present purposes, only the squares marked in roman numerals need to be considered further and these give the principal spurious couplings, which are listed in Table 2. With the formulae for clearances are shown the attenuation requirements, as examples only.

The aim now is to determine values for G, S and IF such that all clearances, K, in Table 2 are adequate. Before this can be done, allowances must be made for frequency modulation effects in both the disturbing and disturbed frequencies, since these effects will reduce the clearances.

(7) NET FREQUENCY CLEARANCES

The disturbing and disturbed frequencies will vary owing to a number of causes. A television channel, for instance, produces considerable frequency deviation under some conditions and such a deviation must be allowed for. These effects are particularly severe in modifying the frequency of the non-linear effect in column E, Table 1, since this depends on twice one channel frequency minus another (or it could depend on the sum of two channel frequencies minus a third). The power in channels which carry f.m. telephony is dispersed more or less according to the speech loading of the channel. Allowance may also be required for inaccuracies in the individual disturbing and disturbed frequencies; when these differ from their nominal values, the effective clearance will again be reduced. The net clearance N for a particular coupling will be the gross clearance, K, less a reduction for the effects just mentioned. Table 3 shows the estimated values for the reductions which are believed suitable for the type of system under discussion, but the present method of treating interference is only a convenient and approximate one.

In a well-proportioned system, the smallest N-value should be as large as possible. Suitable values for S, G and IF may be found by considering the N-values in Table 3 and the gross clearances in Table 2.

(8) OPTIMUM SPACING PARAMETERS

Optimum arrangements can be found by equating three of the first four N-values in Tables 2 and 3 and solving these three

Table 2

PRINCIPAL SPURIOUS COUPLINGS

Type of coupling	Square in Table 1	Description of coupling	Maximum discrimination requirement	Gross clearance formula	Symbol for minimum gross clearance
I	D *a*	Receiver local oscillator to receiver channel.	110 dB	$IF + 4S$ to $IF - 4S$	K_1
I (*a*)	B *e*	Transmitter local oscillator to remote receiver.	92 dB		
II	A *b*	Transmitter channel to LO_1 region.	137 dB	$G - IF + 10S$ to $G - IF + 2S$	K_2
II (*a*)	B *c*	Transmitter local oscillator to image region.	147 dB		
II (*b*)	D *e*	Receiver local oscillator to remote receiver.	55 dB		
III	A *c*	Transmitter channel to image region.	147 dB	$G - 2IF + 10S$ to $G - 2IF + 2S$	K_3
IV	E *a*	Non-linear product to receiver channel.	30 dB	$G + 6S$ to $G - 2S$	K_4
V	C *e*	Unwanted sideband to remote receiver.	92 dB	$2IF + 4S$ to $2IF - 4S$	K_5
V (*a*)	D *d*	Local oscillator to $-IF$ region.	95 dB		

Type IV cannot be dealt with by filtering.

Table 3

NET CLEARANCES

Type I	$N_1 = K_1 - 2$ Mc/s
Type II	$N_2 = K_2 - 3$ Mc/s
Type III	$N_3 = K_3 - 3$ Mc/s
Type IV	$N_4 = K_4 - 7$ Mc/s
Tpye V	$N_5 = K_5 - 3$ Mc/s

equations for S, G and IF in terms of N, while making sure that the remaining equations do not give a lower N-value than the three used for the solution.

A particular solution has been obtained by taking suitable integral values for the multiples of $2S$ in Table 2. It is not difficult to estimate what these values must be for the present problem. The i.f. frequency is fixed at 70 Mc/s and the total bandwidth at about 400 Mc/s, which must allow for 2×6 channels, for G and for the alternative allocation [see Section 2(v)]. The overall bandwidth required is $12 \cdot 5S + G = 400$ Mc/s approximately. G must be near S, $3S$ or $5S$ to get a good clearance for Type IV coupling. S must then be between 25·8 and 29·6 Mc/s. It is now possible to see which multiples of S render critical the different terms in Table 2 and then to solve the equations for clearances of Types I, III and IV.

For present purposes the solution is:

$$S = 2N + 7 \text{ Mc/s}$$
$$G = 3N + 7 \text{ Mc/s}$$
$$IF = 5N + 16 \text{ Mc/s}$$

In these formulae N is the net clearance. If N is specified, as it might well be, the three parameters follow directly. Since the value of IF should preferably be 70 Mc/s, the N-value may be calculated from the third equation and used in the first two. It is then found that

$$N = 10 \cdot 8 \text{ Mc/s}$$
$$S = 28 \cdot 6 \text{ Mc/s}$$
$$G = 39 \cdot 4 \text{ Mc/s}$$
$$IF = 70 \text{ Mc/s}$$

This means that a net clearance of 10·8 Mc/s should be attainable.

It was felt that round numbers would simplify the specifications and that the above figures justified an attempt at such an approximation.

(9) FINAL CHOICE OF FREQUENCY PATTERN

If S had been fixed at 29 Mc/s it would be possible to obtain a net clearance of 11 Mc/s with $IF = 71$ Mc/s and $G = 40$ Mc/s. It was, however, thought desirable to leave the intermediate frequency at 70 Mc/s even at the cost of a slight reduction in net clearance.

The following values were chosen finally:

$$S = 29 \text{ Mc/s}$$
$$G = 39 \text{ Mc/s}$$
$$IF = 70 \text{ Mc/s}$$

These give the following net clearances:

$$N_1 = 10 \text{ Mc/s} \qquad N_3 = 12 \text{ Mc/s} \qquad N_5 = 21 \text{ Mc/s}$$
$$N_2 = 24 \text{ Mc/s} \qquad N_4 = 12 \text{ Mc/s}$$

The worst clearance is slightly smaller than could have been obtained, but the arrangement should permit a baseband range extending to about 10 Mc/s, without giving rise to any of the high discrimination requirements in Table 2. The clearance N_4 is adequate, and this is very important since interference of this type cannot be reduced by providing selectivity.

The overall bandwidth according to the formula is 401·5 Mc/s, but this assumes sidebands of width 14·5 Mc/s beyond the extreme carrier frequencies, where in fact only about 10 Mc/s should be needed ($N = 10$). The true bandwidth is therefore 392·5 Mc/s, and a 'guard band' of 3·75 Mc/s may then be allocated at each end of the pattern to make the total bandwidth equal to the required 400 Mc/s.

Coupling effects also arise between odd- and even-numbered channels. These are much less severe than the effects considered in the paper.

When separate aerials are used, the coupling path between odd and even channels will be from one aerial to the other due to direct aerial coupling as well as to reflections from the landscape in front of the pair of aerials. This coupling will nearly always give a transmission loss which reduces the interference effects to manageable proportions.

The frequency clearances for such couplings can be worked out by extending Table 1. It is only necessary to insert the appropriate odd S-multiples for the disturbing or the disturbed frequency regions.

Couplings between odd and even channels also occur at the remote receiver, and are reduced by the use of relative cross-polarization between these channels. Some of these couplings give rise to the basic selectivity requirement which arises from the finite channel-frequency spacing. The aim of the preceding work has been to choose other parameters such that other, less basic, effects do not lead to any serious increase above the basic selectivity requirements.

This adjacent-channel coupling effect sets a limit to the maximum utilization of the baseband range, and this is a natural and appropriate limitation associated with the channel spacing, S. The spacing of 29 Mc/s actually chosen will not only be sufficient for the present specification (600 telephone channels or a television channel) but will permit a very substantial extension of the baseband range.

(10) CONCLUSION

If the coupling effects between the even (or odd) channels had not been taken into account in choosing the frequency pattern, it is most probable that some selectivity requirements would have been considerably larger than with the chosen parameters, and probably the amplitude non-linearities in the waveguide would have caused interference which would have limited the useful baseband frequency range, or even made it impossible to use a single aerial for transmission and reception.

(11) ACKNOWLEDGMENTS

The author wishes to thank Mr. F. O. Roe for help in preparing the paper for publication, and Standard Telephones and Cables Ltd., for permission to publish it.

(12) REFERENCES

(1) Roetken, A. A., Smith, K. D., and Friis, R. W.: 'The TD-2 Microwave Radio Relay System', *Bell System Technical Journal*, 1951, **30**, p. 1041.

(2) Kaylor, R. L.: 'A Statistical Study of Selective Fading of Super-high Frequency Radio Signals', *ibid.*, 1953, 32, p. 1187 (Fig. 9).

(3) C.C.I.R. VIIIth Plenary Assembly, Warsaw, 1956 (I.T.U., Geneva, 1957), **1**, p. 567.

(4) Jacobsen, B. B.: 'Thermal Noise in Multi-Section Radio Links', *Proceedings I.E.E.*, Monograph No. 262 R, November, 1957 (**105** C, p. 139).

(13) APPENDIX

Estimation of Tolerable Interference in the Main Channel Region (Fig. 6)

In Reference 4 the concept of phase noise was defined (Section 18.1). If in an f.m. (or p.m.) radio link a disturbing frequency is received which is of power $-A$ dB relative to the

nted carrier, this will have the effect of causing phase noise the wanted transmission. The phase noise 'power' will $(-A - 3)$ dB relative to a phase deviation 'power' of ad r.m.s.).[2] This may for convenience be expressed as $A - 3)$ dBrr. This is accurate if A is large (30 or 40 dB). he disturbed and/or disturbing waves are modulated in phase frequency, the expression is still accurate, but the phase noise, tead of having a single-line spectrum, will be widened. This iening is referred to as 'dispersal'. When the radio system ries an f.d.m. system the acceptable amount of interference the baseband is specified for each single telephone channel, ich is about 4 kc/s wide. If the phase noise mentioned above a single-line spectrum all the phase noise will affect one ticular telephone channel, namely that channel which is ated at a baseband frequency equal to the difference between carrier and disturbing wave frequencies. The translation of ase noise into baseband noise depends on the phase modu- on constant of the radio system. In a 600-channel f.m. system ng 6 dB per octave pre-emphasis for the upper half of the seband range, a reasonable value for the phase-deviation nstant for the telephone in the upper baseband range is 15 rad r.m.s. for 0 dBm0 signal or $-18{\cdot}8$ dBrr. This corre- nds to a frequency deviation of 300 kc/s r.m.s. for a signal ut of 1 mW at a point of zero relative level at a baseband quency of 2·6 Mc/s.

An interfering wave of power A dB below the wanted f.m. rier will, if the frequency difference falls in the upper base- nd range, produce noise of value $(-A - 3 + 18{\cdot}8)$ dBm0 a particular channel. If the frequency difference changes, the ise will fall into different telephone channels. Again, if the nted carrier is frequency modulated, the phase noise-fre- ency will similarly be phase modulated and, moreover, with same degree of modulation. The phase noise spectrum will longer be a single-line spectrum; the phase noise may be d to be dispersed. This has the effect of distributing the ase noise over a number of telephone channels and no one nnel will receive as much noise as before modulation was plied. The degree of dispersal depends on the phase deviation the signal and is therefore particularly small for signals in upper part of the baseband range and, of course, small en telephone traffic is light. Signal dispersal cannot, there- re, be relied upon to give much reduction in the noise falling o a single channel. Artificial dispersal has therefore been ed. This consists in adding to the baseband range a low quency, such as 2·8 kc/s for instance, with an amplitude ich gives a frequency deviation of 100 kc/s r.m.s. The corre- nding peak phase deviation is ± 50 rad. A phase deviation this magnitude is very effective in dispersing the phase noise wer. The phase spectrum contains lines at a great many rmonics of the dispersal frequency and no one line contains more than a small fraction of the phase noise power. If both the disturbing and the disturbed waves are artificially dispersed, different frequencies being used, the phase noise is said to be subject to double dispersal. Each line in the spectrum of the single-dispersed signal is then further split into as many lines as before, and if the two dispersal frequencies are suitably dissimilar the phase noise will be rather uniformly distributed over a baseband width of some 400 kc/s and no one 4 kc/s bandwidth will receive more than about 0·01 of the phase noise power, but many such intervals will receive this amount and in fact the total phase noise power remains constant.

In view of the fact that dispersal produces a uniform noise spectrum from an interfering wave, it is possible to relate the effect of such an interfering wave to basic thermal noise. The effect of the interfering wave is equivalent to a worsening of the noise factor of the receiver at sideband frequencies which correspond to the dispersed interfering wave. In the following, the tolerable interference will be calculated on this basis.

The phase noise per 4 kc/s baseband width due to receiver noise is $-P + N - 138$ dBrr, where P is the received carrier signal power in d Bm, N is the noise factor of the receiver and the figure -138 dBm is the noise input power per 4 kc/s under standard conditions.

The total phase noise due to an interfering wave of power I dBm at the input of the receiver is $-P + I - 3$ dBrr. If artificial dispersal is used, the corresponding noise per 4 kc/s will be about 20 dB less or $-P + I - 23$ dBrr per 4 kc/s (over a restricted range of baseband frequencies only). If it is now assumed that the noise factor is 11 dB and that the interference may be allowed to increase the equivalent noise factor by 1 dB it follows that $I = -110$ dBm. This is the figure used in the text for purpose of illustration. It is implicitly assumed that in a long chain of repeaters similar interference will be experienced at each repeater station; if in an actual case only alternate stations receive interference, the permissible level is 3 dB higher.

Similarly, if the interference at successive stations occurs over a range of frequencies, slightly more interference may be allowed per exposure on account of the frequency spread.

If the interference occurs near the carrier frequency of the disturbed channel it will also be possible to allow more noise in so far as thermal noise is not actually a limiting effect for baseband channels in the lowest part of the range in a particular system.

The above shows that, when artificial dispersal is used, the permissible interference can be defined without reference to the received carrier amplitude. This is very convenient, since the received carrier amplitude will be known only after the transmitter power, aerial gain and path length have been determined, and even then it is subject to the variable effects of fading.

SCUSSION BEFORE THE ELECTRONICS AND COMMUNICATIONS SECTION, 7TH DECEMBER, 1959

Mr. W. J. Bray: It is apparent that the preparation of a satis- tory and economical r.f. plan for multi-channel routes is a mplicated matter. A measure of the success of the author's oposals is evident from the fact that the C.C.I.R. has adopted plan, first at the Eighth Plenary Meeting in Warsaw (1956) systems with 600-telephone circuit capacity or a television annel on each r.f. carrier; more recently at the Ninth Plenary eeting in Los Angeles (1959) the plan was again unanimously proved and its use was extended to systems with capacities of 00 telephone circuits or a television channel and up to 600 ephone circuits on each r.f. carrier.

The C.C.I.R. recommendations apply primarily to international nnections, but in fact this plan has importance beyond its application to international connections, since most countries also adopt the recommendations of the C.C.I.R. for their national inland networks. For example, in the United Kingdom the Post Office has installed a number of microwave links based on the author's frequency plan as part of the main trunk network.

The first international application of the plan was, in fact, to the microwave link between the United Kingdom and France, provided jointly by the British Post Office and the French P.T.T. This link provides at the present time four r.f. channels in each direction; one pair, comprising working and standby channels, is used for 600 telephone circuits, and the second pair of channels, each with a bandwidth of 10 Mc/s, is used for 405-, 625- or 819-line television signals.

The plan is economical, not only from the standpoint of equipment design since it enables the amount of r.f. filtration to be kept to a minimum, but also in its use of the r.f. spectrum. In these days of extreme pressure for more space by the many would-be users of the spectrum, this aspect of economy in the use of radio frequencies is most important.

It is apparent that the plan has been derived by an analysis which relates to a particular type of radio repeater, the schematic of which is shown in Fig. 2 of the paper. Were other types of repeater considered in preparing the plan, for example those using all-r.f. amplification, and does the author consider that any serious difficulties may arise in such applications?

Table 3 gives information about the net clearances, and it will be seen that the differences between the gross and net clearances are, in a number of cases, only 2 or 3 Mc/s. This difference corresponds to an allowance for the effective deviations of the wanted and the unwanted signals. If one considers an interfering television signal with a deviation of 8 Mc/s peak-to-peak, it would appear that the difference ought to be at least 4 Mc/s, even without allowing for modulation on the wanted carrier. Thus it would seem that the differences between net and gross clearances shown in Table 3 are rather small, and I would like to ask how the particular values are justified.

In the dispersal scheme described in the Appendix, why was 2·8 kc/s chosen for the modulating frequency? This is an additional modulation imposed on the r.f. carrier and, since practical systems have some non-linearity, there seems a risk of cross-modulation which would result in an interfering 2·8 kc/s signal appearing in the telephone circuits which transmit frequencies up to 3·4 kc/s.

Mr. G. Dawson: In the systems in which this frequency arrangement is in service, a common transmit-receive aerial system is used for up to three two-way channels, and so we were particularly concerned with the amplitude non-linearity effects discussed in Section 5.1; as each system was installed these effects were measured and in almost every case we found that the resultant distortion terms were well below the safe levels quoted in the paper. However, in one or two cases some feeders did produce intermittent excessive distortion, which was traced to mechanical distortion of the waveguide joints. When these faulty joints were replaced the combination terms dropped to a safe level, and this experience has emphasized the importance of ensuring that the feeder has a mechanically sound construction as well as giving good electrical performance.

A point which could have been mentioned more clearly in the paper is the production of fourth-order sidebands in the medium-level mixer. These sidebands fall in a sensitive frequency region in one of the receivers, and since they are generated at a level which is comparable with that of the main output at the wanted frequency of the system, large attenuation is required to obviate any adverse effects. We have found that in some equipments the suppression of this fourth-order sideband has to be taken into account in the design of the output filter, as well as the rejection of the local oscillator of the unwanted sidebands.

Are the 1 800 channels, advocated by the C.C.I.R., the upper limit on this frequency plan or can it be extended even further?

Mr. G. W. S. Griffith: The author was faced with a very difficult job in working out this frequency plan because too many points were specified in advance. It is very difficult to get six 600-channel systems for each direction of transmission into a 400 Mc/s bandwidth using an intermediate frequency of 70 Mc/s. He was not able to decide where he would like his unwanted terms to fall, and he has cleverly devised the best that will go into the space available.

In item (vii) in Section 2 the author points out that one of the important points about his frequency plan is that it allows three transmitters and three receivers to be put on a single transmissio line, and that the interleaved channels (the even-numbere channels if the first three are odd numbers) should use a separa aerial system to avoid some interferences. It has been suggeste in another context that the aerial system of the future might be horn paraboloid fed by a circular waveguide which will transm signals of both polarizations, the size of the waveguide beir chosen to handle more than one frequency band. This entai transmission of signals of both polarizations in the same wav guide. There will be polarization separation, but I do not thir it is fair to design on a figure of more than about 25 dB, and the might well be some cross-coupling in the non-linearity effects.

Also, if this type of aerial is used for the transmission multiple-frequency bands, non-linearity effects may give inte modulation of signals in harmonically related bands. Thus, signal in the 2 000 Mc/s band and one in the 4 000 Mc/s ban may produce a signal on a wanted frequency. This suggest that the centre frequencies of these bands should have bee chosen to minimize this effect; but these frequencies were picke to avoid harmonics of the shift oscillator which, as the autho says, can very often be avoided in the equipment design.

Can the author indicate the levels these interferences migh reach and state whether these problems will set limitations on th use of multiple-band aerials?

Mr. G. Millington: Are the radio links discussed by the autho always optical or do they ever include receivers that are belo the horizon or in the neighbourhood of hills or man-mad obstacles, and in this connection, what protection does he expe to get in decibels from the use of cross-polarization?

In the slide showing a mast on top of a building it looks as the axes of the backward- and forward-looking sets of dishes a not in line. Is this just coincidental or is use being made of th fact that it may be desirable to zigzag the route slightly to reduc the possibility of jump-over from one link to another furth along the chain and working on the same or a very near frequency

With regard to the use of this system in various parts of th world, I should imagine that there may be difficulties in mou tainous terrain or where anomalous propagation may be prev lent at certain times. I should like to know how the syste stands up to such extreme conditions.

The author mentioned the possible trouble that could ari from an obstacle such as a tree near to a terminal. Is any suc disturbance ever caused by reflections from aircraft?

Mr. W. Grossett: The use of dispersal for reducing an inte ference signal is particularly advantageous when using lightl loaded circuits, but its advantages are probably limited whe the system is being worked near its full capacity. With wh order of numbers of channels are the advantages of dispersal t be gained? It is of interest to note that, for certain radio systen in current use, where a supervisory channel is positioned belo the baseband, dispersal is achieved by the use of tones which a used for the transmission of various signalling information.

In general, little consideration has been given to the effects interference associated with television. I think this is becaus the requirements are not very stringent, but there are two facet The recent introduction of a pre-emphasis characteristic ass ciated with a television channel has resulted in the signal bein more susceptible to the effects of interference in the low-frequenc portion of the video-frequency band. This, of course, is th part of the band where the picture is most susceptible to patter interference. However, I would suggest that the disadvanta associated with the pre-emphasis characteristic is of a ver minor nature compared with the advantages to be achieved b pre-emphasis. With the possible advent of colour televisio following the N.T.S.C. form, a more stringent requirement wi exist, particularly with interference signals occurring in the neigh

ırhood of the colour sub-carrier, where probably protection the order of 20 dB has to be achieved as compared with the sent systems.

The author refers to the effects occurring outside the base-nd range in connection with the permitted levels of i.f. inter-ence, where he quotes that the interference shall be below 75 dBm. Can he explain the mechanisms producing such eband noise? One obviously to be considered is the effect signals appearing inside the i.f. band which would catch the omatic gain control system. This would have to be avoided cases of fading.

Mr. J. A. Ratcliffe: May I ask a very elementary question which curs to somebody not working in this field of engineering at ? Why does one get non-linear terms in waveguides, what is ir cause and what has been done to eradicate their effects?

Mr. R. G. Medhurst (*communicated*): The authors' analysis the problem could be given added precision by employing a re exact treatment of the mechanism of interfering carrier tortion. He states that the paper is not concerned with the culation of the permissible amount of spurious signal, and, of urse, for the present purpose qualitative arguments are ade-ate. However, the assumption, as in Section 4, that the width the sensitive regions shown in Fig. 6 is only slightly more than ce the maximum baseband frequency (i.e. ±2·5 Mc/s for a)-channel telephone system) is a drastic oversimplification. was shown some time ago* that the central region of high sitivity extends over twice this range (±5 Mc/s for the 600-nnel system), with a not very rapid fall-off outside. The umption made in deriving this result was that the wanted rier is modulated with the C.C.I.R. recommended white-se test signal, i.e. that the system is operating with average ding.

n the 600-channel case, the clearances shown in Section 9 pear to be adequate. However, the recent C.C.I.R. confer-ce† has adopted the same scheme for up to 1 800 channels. is would appear very hazardous. Fig. A shows interfering rier distortion in the 1 800-channel case. The curve is for an nodulated interfering carrier and neglects the effect of pre-phasis. Modulation of the interfering carrier makes no stantial difference to the shape of the curve, and pre-emphasis

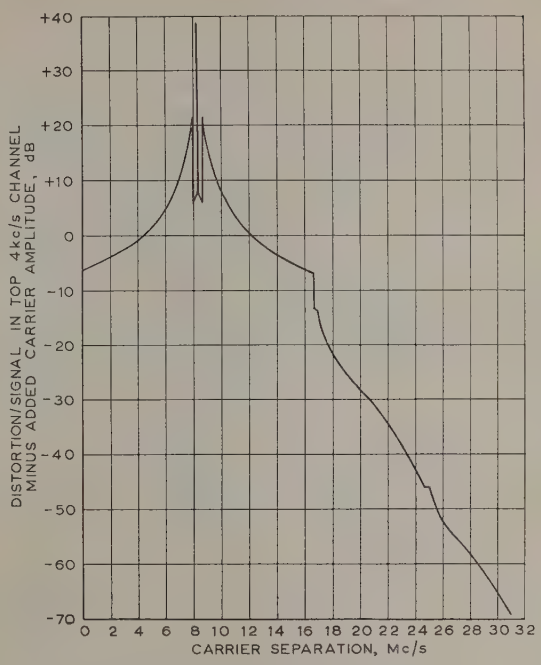


Fig. A.—Distortion due to an interfering carrier in an 1 800-channel system (added carrier unmodulated).

will merely depress it somewhat in the region around 8 Mc/s. The sensitive region extends over ±16 Mc/s and beyond, whereas, from Section 9, clearances (in the author's terminology, gross clearances) of 12 and 15 Mc/s are possible under the present scheme. It would appear, in consequence, that for 1 800-channel systems there is no special virtue in the frequency pattern proposed in the paper, and, in fact, in the frequency space available it would seem very difficult to avoid the danger of excessive vulnerability to interference.

THE AUTHOR'S REPLY TO THE ABOVE DISCUSSION

Mr. B. B. Jacobsen (*in reply*): Some consideration was given the repeater using all-r.f. amplification (single-mixer type) ntioned by Mr. Bray. Some forms of this repeater are subject serious interference when a harmonic of the local-oscillator ift) frequency falls near a channel frequency and it would then necessary to avoid the use of certain shift frequencies for tain channels, an undesirable restriction on planning. Too ny points were already specified before the frequency pattern s designed and there was little or no scope left for introducing ther conditions. When using the single-mixer type of repeater mixer design should, therefore, be one which does not cause due generation of harmonics of the shift frequency.

The dispersal frequency produces some non-linear distortion ect, but it does not appear as a discrete frequency in any of the gle-sideband telephone channels. The chosen frequency is not que; other frequencies are also useful.

Mr. Dawson points out that fourth-order sideband in the dium-level mixer falls very close to the associated receive nnel. It is true that this creates a potential source of inter-ference. The single-sideband filter may, for this reason, need to be designed to provide slightly more selectivity than would otherwise be necessary, but this is not difficult because the frequency to be attenuated is far removed from the pass-band of the filter.

Mr. Griffith raises the problem of non-linearity effects when multiple-frequency bands share an aerial. It is expected that these effects would be comparatively small in the large circular waveguide which would be used. The worst effect may be from channels in the 2 000 Mc/s band on those in the 6 000 Mc/s band. It is usually possible in fully equipped stations to arrange that the channels sharing an aerial give reasonable clearance; e.g. the odd-numbered channels of a particular allocation at 2 000 Mc/s will be compatible with either the odd- or the even-numbered channels of the 6 000 Mc/s systems. When a common send-receive aerial is used it is further possible to provide mutual cross-polarization between the two frequency bands and this may in itself be sufficient to make other restrictions unnecessary.

In reply to Mr. Millington, the radio paths used for these systems are designed to provide line-of-sight transmission and to have a good clearance over the intervening terrain. The cross-polar conversion loss may then be expected safely to exceed 20 dB. It is often necessary to zigzag the route purposely to

* Medhurst, R. G., Hicks, Mrs. E. M., and Grossett, W.: 'Distortion in Frequency-Division-Multiplex F.M. Systems due to an Interfering Carrier', *Proceedings I.E.E.*, Paper No. 2565 R, May, 1958 (105 B, p. 282).

† C.C.I.R. Documents of the IXth Plenary Assembly, Los Angeles, Vol. 1, Recommendations, p. 243 (International Telecommunications Union, Geneva, 1959).

reduce the incidence of 'over-shoot' transmission, i.e. transmission of the same wave directly over three sections, but in some cases natural obstructions give sufficient protection. Transmission systems in the 4 000 Mc/s band are giving good service in many parts of the world because the paths have been carefully planned and the equipment has been designed to accept fading of about 30 dB over any one section without giving too much noise in the telephone channels. Experience in mountainous country has in some cases been particularly favourable. Serious disturbances from aircraft are very rare, but it would not be advisable to install a repeater station near an airfield. The greatest disturbance occurs when an aircraft is in the radiated beam and near one of the aerials.

Mr. Grossett is right in saying that a measure of dispersal is obtained from the signal itself and from pilot frequencies, but the effect is very small since the signal, even in a fully equipped system, may be of very low level outside the peak traffic period. A useful degree of dispersal of the interfering spectrum requires a substantial phase deviation (tens of radians), and this can be attained only by using very low frequencies. Dispersal is beneficial whenever a disturbance falls within a baseband width of the carrier frequency, except when the interfering carrier is very near the wanted signal.

Interference into and from radio channels carrying television was considered from the beginning, and it was very soon found that pre-emphasis of the order now adopted by the C.C.I.R. was desirable, to limit the r.f. spectrum to a reasonable bandwidth.

One mechanism that leads to the –75 dBm requirement mentioned by Mr. Grossett is the following: in one repeater station an interfering frequency, say 12 Mc/s from the carrier, will be incorporated in the f.m. signal by the limiter. If, in a subsequent station, there is similar interference at, say, 15 Mc/s, then after limiting, the two effects together will produce interference centred on a baseband frequency of 3 Mc/s and with 'phase power' $A + B - 9$ dBrr, where A and B are the margins between the interference and the unmodulated carrier. Still furth interfering frequencies will combine with each of the origin effects in a similar way and there may thus be a great ma interference terms in the baseband.

In reply to Mr. Radcliffe, the non-linear effects occur whe wall currents flow through imperfect joints in the waveguid e.g. at a flange, or where there is a high current density (resona structure) and the surface carrying the current is impure, e. after soldering. Junctions of this nature will certainly tend give third-order distortion, but it is not known if second-ord distortion also occurs; it seems less likely on account of t symmetry of many imperfections with respect to the direction current flow. A firm, purely metallic contact produces le distortion than can be measured, but a low-pressure contact an impure surface can produce even more severe distortion tha that mentioned in the paper.

Mr. Medhurst's objections to the assumptions in Section have been anticipated in that Section, where it is said that a extra allowance is needed when the disturbed and/or disturbir frequency is modulated, but I did not state explicitly that th allowances in Table 3 are those which apply for the 600-chann system which was considered when the frequency plan w studied. These allowances were, of course, based on the f.n spectrum of the signals.

The slope of Mr. Medhurst's curve agrees accurately wit calculations made for an 1 800-channel system with pre-emphasi but near and below 8·2 Mc/s the shape for the system with pr emphasis is distinctly different due to dispersal and the whol curve is lower and smoother. In an 1 800-channel system th gross clearance of 12 Mc/s is not as large as could have bee desired but is still better than the 6 Mc/s clearance provided b the previous plan. I agree with Mr. Medhurst that in an 1 80 channel system it is difficult to avoid vulnerability, but then th cost of providing the necessary selectivity is shared by a large number of telephone channels.

.372.553 : 621.397.2

The Institution of Electrical Engineers
Paper No. 3054 E
Oct. 1959

A QUADRATURE NETWORK FOR GENERATING VESTIGIAL-SIDEBAND SIGNALS

By G. G. GOURIET, Member, and G. F. NEWELL, Associate Member.

e paper was first received 11th February, and in revised form 3rd June, 1959. It was published in October, 1959, and was read before the ELECTRONICS AND COMMUNICATIONS SECTION *11th January, 1960.)*

SUMMARY

he paper describes a method of generating approximately a idrature signal from a given signal using a linear network com- sing a tapped delay line. It is shown that a vestigial-sideband signal e from phase distortion can be produced by this means.

An experimental version of the quadrature network described has n constructed, and the paper includes the results obtained using the work to produce a vestigial-sideband version of a 405-line television nal with a video bandwidth of 3 Mc/s.

t is suggested that a high-power vestigial-sideband signal could be oduced by this means with some saving in transmitter power.

LIST OF PRINCIPAL SYMBOLS

$s(t)$ = Arbitrary function of time.
$\hat{s}(t)$ = Hilbert transform of $s(t)$.
t = Time, sec.
T = Time interval, sec.
τ_0 = Relative time delay of adjacent taps on a line.
$\delta(t)$ = Dirac function or unit impulse.
$u(\omega)$ = Heaviside unit step occuring along the ω-axis at $\omega = 0$.
v], $s^*[\omega]$, $s[\omega, T]$, $H[t]$ = Fourier transforms of $s(t)$, $s^*(t)$, $s(t, T)$ and $H(\omega)$, respectively.
ω = Angular velocity in rad/s.
Δ = Numerical fractions.
α = Design parameter of an amplitude/frequency characteristic.
n = Number of sections forming a delay line.
r = Tap number in order from the line centre.
A = Constant.
θ = Arbitrary phase angle.

(1) INTRODUCTION

Methods of generating single-sideband or vestigial-sideband nals fall into two main groups,[1] one of which requires the neration of a double-sideband signal, with or without a carrier, d the subsequent removal of the unwanted sideband by filters. ie other method,[2] generally called the phase-shift method, tails the generation of two double-sideband signals which, ien added together, form a single or vestigial-sideband signal. iere is a so-called 'third method',[3] which might be considered a combination of the other two methods, in that it relies on a :er to remove the unwanted sideband components near to the inted sideband.

The design of a filter capable of dissipating high power and removing the unwanted sideband, without introducing appre- ible distortion in the wanted sideband, presents some difficulty. television broadcasting this problem is partly avoided by ing a form of asymmetric-sideband characteristic which places e onus of attempting dispersion-free filtering on the receiver signer.

The phase-shift method removes the necessity for filtering but requires the ability to change, by $\pi/2$ rad, the phase of the carrier and all the spectral components of the modulating signal. Fig. 1 shows the schematic of the phase-shift method and typical spectra produced.

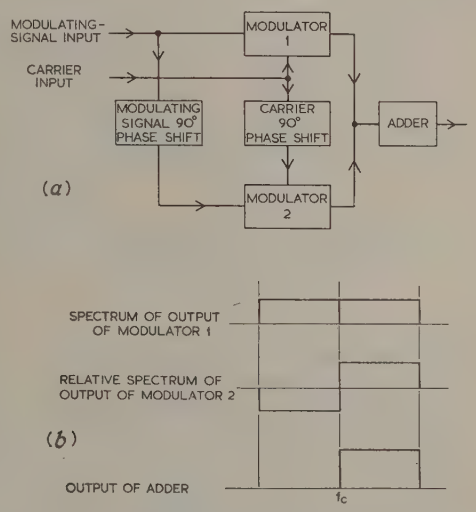


Fig. 1.—Phase-shift method of single-sideband generation.
(*a*) Schematic arrangement.
(*b*) Resulting spectra.

(1.1) Simple Phase-Shift Networks

The circuits in Fig. 2 are capable of producing approximately a $\pi/2$ phase shift for frequencies either much higher [Figs. 2(*a*) and (*b*)] or much lower [Figs. 2(*c*) and (*d*)] than a chosen design frequency. Unfortunately, such circuits all produce distortion

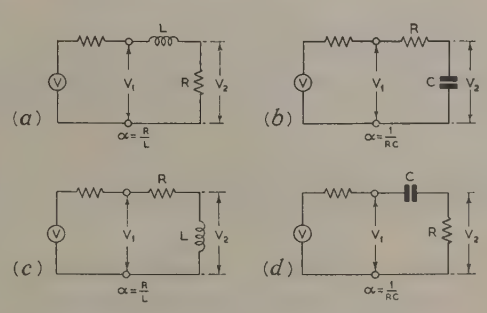


Fig. 2.—Elementary phase-shifting networks.

(*a*) and (*b*) Low-pass circuits: $V_2 = V_1\sqrt{\left(\frac{\alpha^2}{\alpha^2+\omega^2}\right)} \angle \arctan \omega/\alpha.$

(*c*) and (*d*) High-pass circuits: $V_2 = V_1\sqrt{\left(\frac{\omega^2}{\alpha^2+\omega^2}\right)} \angle \arctan -\alpha/\omega.$

Mr. Gouriet was formerly, and Mr. Newell is, in the Research Department of the tish Broadcasting Corporation. Mr. Gouriet is now with the Wayne Kerr oratories Ltd.

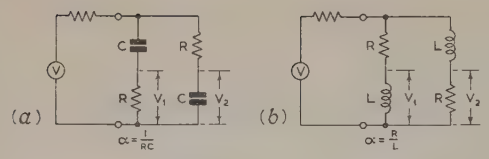


Fig. 3.—Six-terminal networks providing two outputs V_1 and V_2.

$$V_2 = \frac{V_1\omega}{\alpha} \angle \pi/2.$$

of the amplitude/frequency characteristic. The networks shown in Fig. 3 produce a $\pi/2$ phase difference at all frequencies but have an amplitude response proportional to frequency. Fig. 4(*a*) shows a similar circuit arranged as an all-pass network having a phase relationship which is a function of frequency, as

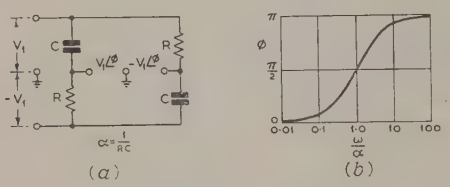


Fig. 4.—Simple all-pass network.
(*a*) Circuit diagram.
(*b*) Phase/frequency characteristic.

shown in Fig. 4(*b*). Two such networks with slightly different values of the design parameter, α, will produce two outputs differing only by their relative phase angle, which can be kept approximately constant over a band of frequencies. The use of several such pairs of networks in cascade can produce a relative phase angle of $\pi/2$ rad over a band of frequencies. All-pass networks[4–7] enable the phase-shift method to be used successfully for telephony and even for high-quality audio transmission, but they all introduce dispersion at frequencies above and below the passband. Whilst this dispersion is not serious for audio transmission, the low-frequency dispersion would mar television quality. Fig. 5 shows the arrangement of such a system and a typical phase characteristic. In the following Sections it is intended to describe a network which produces a $\pi/2$ phase shift at all frequencies below a chosen design frequency and which has certain advantages over all others described so far for vestigial-sideband transmission.

(2) AN APPROXIMATE REALIZATION OF 'QUADRATURE'

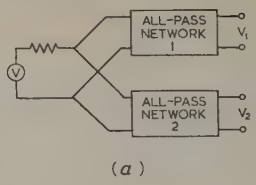
The definition and significance of quadrature signals have been discussed by Gabor,[8] who has suggested a mechanical means of generating such signals. This paper is concerned primarily with an alternative method which employs a 4-terminal linear electrical network.

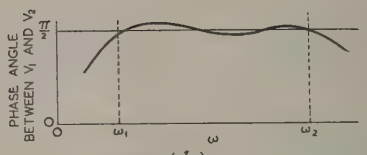
By definition, two signals, s(t) and s*(t), are in quadrature if they are related by the pair of Hilbert transforms:

$$s^*(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{s(\tau)}{t-\tau}d\tau \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$s(t) = -\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{s^*(\tau)}{t-\tau}d\tau \quad . \quad . \quad . \quad . \quad (2)$$

The term 'quadrature' arises because of the significance of the Hilbert transform in terms of spectra; the equivalent operation





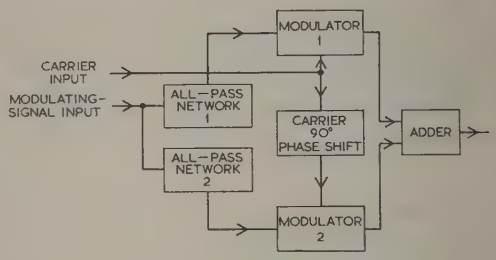


Fig. 5.—All-pass networks.
(*a*) Six-terminal network comprising two all-pass networks.
$|V_1| = |V_2| = |V|$
(*b*) Typical phase relationship between V_1 and V_2.
(*c*) Schematic arrangement using all-pass networks to generate a single-sideban signal.

is to change the phase of every Fourier component by $\pi/2$, an amounts to multiplying the complex spectrum by the imaginar operator j in the negative frequency domain and by $-j$ in th positive domain.

In seeking a network which will perform this operation it i pertinent to inquire what its response would be to a Dira function or unit impulse defined by the Fourier integral as

$$\delta(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \varepsilon^{j\omega t}d\omega$$

If in eqn. (1) s(τ) is replaced by $\delta(\tau)$, we obtain

$$\delta^*(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\delta(\tau)}{t-\tau}d\tau = \frac{1}{\pi t} \quad . \quad . \quad . \quad .$$

by virtue of the properties of the Dirac function $\delta(t)$.

The equivalent operation in terms of the spectrum is confirme by comparing the spectra $\delta[\omega]$ and $\delta^*[\omega]$.†

For the spectrum of $\delta^*(t)$ we have

$$\delta^*[\omega] = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\varepsilon^{-j\omega t}}{t}dt = -j \operatorname{sgn}\omega \quad . \quad . \quad .$$

where sgn ω, the solution of Dirichlet's integral, is defined by

$$\frac{2}{\pi}\int_0^{\infty} \frac{\sin \omega t}{t}dt = \operatorname{sgn}\omega = \begin{cases} 1 & \omega > 0 \\ 0 & \omega = 0 \\ -1 & \omega < 0 \end{cases}$$

Comparison of eqn. (4) with the spectrum of $\delta(t)$, which unity for all values of ω, confirms the previous interpretatio

† A convention suggested by Dr. R. D. A. Maurice is used here in which t Fourier transform of f(t) is written f[ω].

t follows that the spectrum of $\frac{1}{2}[\delta(t) + j\delta^*(t)]$ is

$$\left.\begin{array}{ll} 1 & \omega > 0 \\ \frac{1}{2} & \omega = 0 \\ 0 & \omega < 0 \end{array}\right\} = \mathrm{H}(\omega) \quad . \quad . \quad . \quad . \quad (5)$$

where $\mathrm{H}(\omega)$ is Heaviside's unit function.

In general, if $s(t)$ has a spectrum $s[\omega]$, the complex time unction

$$\psi(t) = \tfrac{1}{2}[s(t) + js^*(t)]$$

vill have a spectrum

$$\psi[\omega] = s[\omega]\mathrm{H}(\omega) \quad . \quad . \quad . \quad . \quad . \quad (6)$$

nd it is this property which may in theory be used to produce a ingle-sideband carrier signal without the conventional filtering rocess.

It is known that a perfect quadrature signal cannot be chieved. It is, however, possible to devise a linear network vhich may be used to approximate closely to such a signal over finite bandwidth, and a vestigial-sideband radio-frequency ignal can be produced by this means with advantages over con-entional methods.

(3) THE QUADRATURE NETWORK

It follows from what has been said that a linear 4-terminal etwork with an impulse response, $1/t$, will produce a quadrature ignal $s^*(t)$ when (t) is applied to its input. Such a network nay be realized to a required degree of approximation by using delay line in the manner used by Linke[9] for the purpose of qualization.

Consider a terminated lumped delay line tapped at regular ntervals along its length, and with unit impulse applied to the nput. If the output at each tap is adjusted in amplitude and ign to be inversely proportional to the time delay relative to the entre of the line (see Fig. 6), the sum of the outputs at all taps

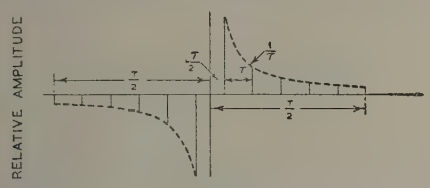


Fig. 6.—Amplitude distribution at taps along line.

vill (apart from a fixed delay) tend to $1/t$ as the length of the line ends to infinity and the interval between taps tends to zero. or a finite interval, τ_0, between the taps the output in the requency band $0 - 1/2\tau_0$ cycles per second will correspond to he signal $1/t$ with a similarly restricted bandwidth. Thus, ssuming for the moment an infinite line, the response to unit mpulse in the band 0–ω_0 is given by

$$s^*(t) = \frac{1}{2\pi j}\int_{-\omega_0}^{+\omega_0} \varepsilon^{j\omega t}\,\mathrm{sgn}\,\omega d\omega$$

$$= \frac{1}{\pi}\int_0^{\omega_0} \sin \omega t d\omega = \frac{1}{\pi t}(1 - \cos \omega_0 t) \quad . \quad . \quad . \quad (7)$$

$^*(t)$ in eqn. (7) is thus the quadrature signal derived from the gnal

$$s(t) = \frac{1}{2\pi}\int_{-\omega_0}^{\omega_0} \varepsilon^{j\omega t} d\omega = \frac{\sin \omega_0 t}{\pi t} \quad . \quad . \quad . \quad . \quad (8)$$

nd is shown plotted in Fig. 7(*a*).

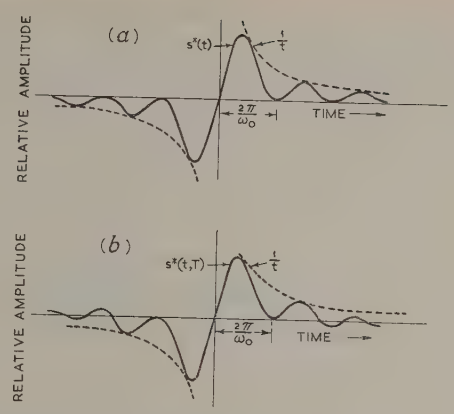


Fig. 7.—Quadrature signal $s^*(t)$ derived from eqns. (8) and (12).
(*a*) Quadrature signal $s^*(t)$ corresponding to $s(t) = \sin(\omega_0 t)/\pi t$.
(*b*) Approximate quadrature signal produced with a line of finite length T, computed from eqn. (12).

A further limitation arises in practice owing to the finite length of the line. Clearly this will have the greatest effect at low frequencies, and the modified spectrum due to this cause can be derived as follows:

Assuming the total length of the line to be T, the response to unit impulse apart from a delay $T/2$ will be 0 for $t < -T/2$, $1/t$ for $-T/2 < t < T/2$ and 0 for $t > T/2$, and eqn. (4) will become

$$\delta^*[\omega, T] = \frac{1}{\pi}\int_{-T/2}^{T/2} \frac{\varepsilon^{-j\omega t}}{t} dt = \frac{2}{j\pi}\,\mathrm{Si}\left(\frac{\omega T}{2}\right) \quad . \quad . \quad (9)$$

To evaluate the response $s^*(t, T)$ of the finite line to unit impulse in the band 0–$\omega_0/2\pi$, we may write

$$\frac{\pi}{2}s^*(t, T) = \frac{1}{2\pi}\int_{-\omega_0}^{\omega_0} \mathrm{Si}\left(\frac{\omega T}{2}\right)\sin \omega t d\omega \quad . \quad . \quad (10)$$

or, alternatively, treat the problem in terms of the convolution integral and write

$$s^*(t, T) = \int_{t-T/2}^{t+T/2} \left(\frac{\sin \omega_0 \tau}{\pi\tau}\right)\left[\frac{1}{\pi(t-\tau)}\right] d\tau \quad . \quad . \quad (11)$$

Either of these integrals gives for the solution

$$s^*(t, T) = \frac{1}{\pi^2 t}\left\{\mathrm{Si}\left[\omega_0\left(t + \frac{T}{2}\right)\right] - \mathrm{Si}\left[\omega_0\left(t - \frac{T}{2}\right)\right] - 2\cos \omega_0 t\,\mathrm{Si}\left(\omega_0\frac{T}{2}\right)\right\} \quad . \quad (12)$$

which as T tends to infinity gives the previous result:

$$s^*(t) = \frac{1}{\pi t}(1 - \cos \omega_0 t)$$

Eqn. (12) is shown plotted in Fig. 7(*b*) using the value $\omega_0 T/2 = 3 \cdot 6\pi$.

(4) VESTIGIAL-SIDEBAND SIGNAL GENERATION

It has been shown that a complex signal of the form

$$\psi(t) = \tfrac{1}{2}[s(t) + js^*(t)]$$

has a spectrum

$$\psi[\omega] = s[\omega]\mathrm{H}(\omega) \quad . \quad . \quad . \quad . \quad . \quad (13)$$

and is therefore a 'single-sideband' signal with respect to zero frequency. The imaginary time function cannot, of course, be realized physically except as an orthogonal component in a 2-dimensional field. As pointed out by Gabor, however, given $s^*(t)$ it is possible to produce the same single-sideband spectrum but with the origin moved to a frequency $\omega_c/2\pi$. Thus

$$\psi(t, \omega_c) = \tfrac{1}{2}[s(t) \cos \omega_c t \pm s^*(t) \sin \omega_c t] \quad . \quad . \quad (14)$$

has a spectrum

$$\psi[\omega, \omega_c] = s[\pm(\omega_c - \omega)]H(\pm\langle\omega_c - \omega\rangle) \quad . \quad . \quad (15)$$

for which the upper or lower sideband with respect to ω_c is suppressed, depending on whether the sign in eqn. (14) is positive or negative.

By substituting in eqn. (14)

$$s(t) = \frac{\sin \omega_0 t}{\pi t}$$

and

$$s^*(t) = \frac{1}{\pi t}(1 - \cos \omega_0 t)$$

it is easy to confirm that the result in the positive-frequency domain is a uniform spectrum extending either from ω_c to $(\omega_c - \omega_0)$ or $(\omega_c + \omega_0)$, depending on whether the quadrature component is added or subtracted.

Owing to the practical limitations of the quadrature network, as discussed, the signal $s^*[t, T]$ given in eqn. (12) must be substituted for $s^*(t)$, and the resulting vestigial-sideband spectra obtained in practice are shown in Fig. 8. The ripples, which are

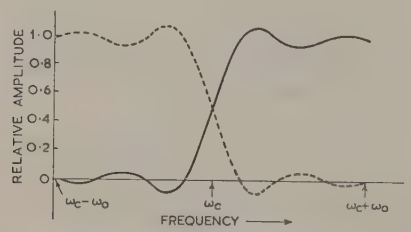


Fig. 8.—Vestigial-sideband spectra obtained with approximate quadrature signal from finite line.

inherently due to the finite length of the line, cannot be reduced by any finite increase in the length: eqn. (9) shows that the effect of increasing T is simply to increase the frequency of the ripples. A monotonic spectrum may, however, be approximately achieved with a finite line if the 'tapering' function $1/\tau$ is modified to approach zero at a suitably faster rate. For example, if in eqn. (9) $1/t$ is replaced by $\varepsilon^{-a|t|}/t$ and the limits are made infinite to correspond to an infinite line, we obtain, for the modified spectrum,

$$\phi[\omega] = \int_{-\infty}^{\infty} \frac{\varepsilon^{-a|t|}}{\pi t} \varepsilon^{-j\omega t} dt$$

$$= \frac{2}{j\pi} \text{ arc tan } \frac{\omega}{\alpha} \quad . \quad . \quad . \quad . \quad . \quad (16)$$

which is monotonic.

If the line is now made a finite length, T', which is so chosen that

$$\varepsilon^{-\alpha T'/2} \ll 1 \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

the response will be substantially unchanged, since the contributions at the extremities will be negligible. Thus, the value of α is chosen to give the required frequency response according to eqn. (16) whilst the length of the line is chosen to satisfy eqn. (17).

A somewhat better alternative is to replace $1/\pi t$ in eqn. (4) by the function

$$f(t) = \frac{\alpha}{2} \text{ cosech } \frac{\pi\alpha t}{2}$$

This provides a monotonic spectrum

$$f[\omega] = \tanh \frac{\omega}{\alpha}$$

which has the advantage of a relatively fast rate of cut-off in the low-frequency region.

(5) PRACTICAL QUADRATURE NETWORK

In practice the delay line is tapped at finite intervals, and this has an important effect on the frequency response of the network. Sampling a time function at regularly spaced discrete times produces a frequency response which is a periodic function of frequency. By arranging the taps so that the discontinuity at time $t = 0$ falls between two taps, the resulting frequency response is as shown in Fig. 9. The spectrum changes sign at

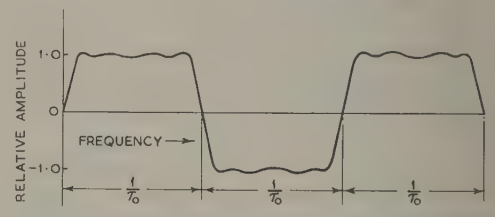


Fig. 9.—Frequency response of a quadrature network having a finite length of ideal line tapped at regular intervals.

frequency spacings equal to $1/\tau_0$,where τ_0 is the time interval between taps. The frequency band occupied by one complete cycle of the spectrum is $1/0{\cdot}5\tau_0$; the time interval between the centre of the line and the nearest tap is $0{\cdot}5\tau_0$. This, of course, is completely analogous to the more familiar sampled frequency spectrum producing a periodic continuous time function.

The effective bandwidth of the practical quadrature network will, in general, be limited by the bandwidth of the delay line used. In the case of a lumped-circuit delay line, it would be uneconomic to make the bandwidth greater than the reciprocal of the time interval between taps.

The practical network consists of an odd number of sections of delay line forming a continuous terminated line. The line has a voltage tap at each junction, and the output of those taps between the input and the centre are added and polarity reversed before being added to the outputs from the remaining taps. The amplitude of the outputs of the individual taps must be adjusted according to the relationship:

$$\text{Output amplitude at each tap} = \frac{\text{cosech } \frac{1}{4}\pi r\alpha\tau_0}{\text{cosech } \frac{1}{4}\pi\alpha\tau_0} \quad . \quad (18)$$

where r = Number of half-sections between the tap and the centre of the line (negative for taps between the input and the centre).

τ_0 = Delay between adjacent taps.

α = Design parameter determining the slope $(= 1/\alpha)$ of the amplitude/frequency characteristic at zero modulation frequency.

For an input signal $s_1(t)$ the output of this network is given by

$$s_2(t) = A \sum_{x=-(n-1)/2}^{(n+1)/2} s_1\left[t - \frac{n\tau_0}{2} + \frac{(2x-1)\tau_0}{2}\right] \operatorname{cosech} \frac{\pi\alpha\tau_0(2x-1)}{4} \quad . \quad . \quad . \quad . \quad (19)$$

where n is the total number of sections in the line.

If we consider the case where $s_1(t) = \cos(\omega t + \phi)$, the network output is given by

$$s_2(t) = -\sin\left[\omega\left(t - \frac{n\tau_0}{2}\right) + \phi\right] A \sum_{x=-(n-1)/2}^{(n+1)/2} \sin \frac{(2x-1)\omega\tau_0}{2} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x-1)}{4} + \cos\left[\omega\left(t - \frac{n\tau_0}{2}\right) + \phi\right] A \sum_{x=-(n-1)/2}^{(n+1)/2} \cos \frac{(2x-1)\omega\tau_0}{2} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x-1)}{4} \quad . \quad (20)$$

in which the second term sums to zero.

Provision of a separate centre tap on the line makes available the input signal delayed in time by half the total delay of the line. This delayed input signal can be written

$$s_1(t, \tau) = \cos\left[\omega\left(t - \frac{n\tau_0}{2}\right) + \phi\right] \quad . \quad . \quad . \quad (21)$$

Therefore, choosing the time reference to be that of $s_1(t, \tau)$, we can rewrite the first term of eqn. (20) as

$$s_2(t) = s_1(t, \tau) A \sum_{x=-(n-1)/2}^{(n+1)/2} \sin \frac{(2x-1)\omega\tau_0}{2} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x-1)}{4} \quad . \quad . \quad . \quad . \quad (22)$$

By virtue of superposition it follows that the output of the quadrature network will consist of the sum of all the frequency components of $s_1(t, \tau)$ each rotated in phase by $-\pi/2$ rad and modified in amplitude by the term $f[\omega]$, where

$$f[\omega] = A \sum_{x=-(n-1)/2}^{(n+1)/2} \sin \frac{(2x-1)\omega\tau_0}{2} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x-1)}{4} \quad . \quad (23)$$

In Fig. 10, $f[\omega]$ is plotted as a function of normalized frequency $\omega\tau_0/2\pi$ for three cases of a 15-section network with $\pi\alpha\tau_0$ equal to 1·6, 1·0 and 0·6, respectively. In each case A was made equal to $(2/\pi)\sinh\frac{1}{4}\pi\alpha\tau_0$ so that the output at the taps nearest the centre of the line would be the same for all three cases. Fig. 10 shows that the value of $\pi\alpha\tau_0$ does not materially affect the mean amplitude of the responses at the middle frequencies, but affects the amplitude of the ripples and the slope of the curve at the frequency limits $\omega\tau_0/2\pi = 0$ and 1·0. The larger the value of $\pi\alpha\tau_0$, the smaller the ripple amplitude and the slope at the frequency limits. The curve for $\pi\alpha\tau_0 = 1{\cdot}6$ has a ripple amplitude less than 1·0% and has an effective normalized bandwidth equal to 0·75 (frequency limit $\simeq 0{\cdot}75/\tau_0$).

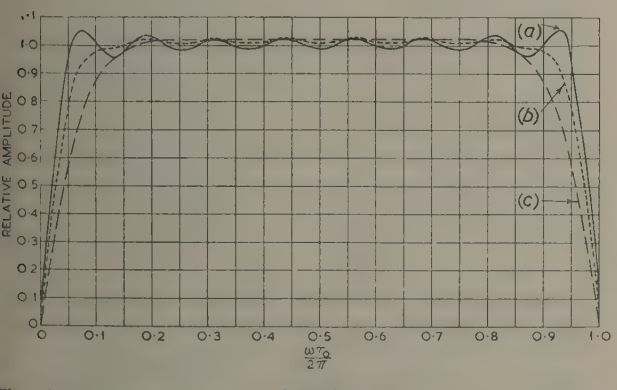


Fig. 10.—Frequency responses of quadrature network having 15 sections of line.

(a) ——— $\alpha\tau_0 = 0{\cdot}6/\pi$.
(b) – – – $\alpha\tau_0 = 1/\pi$.
(c) — — — $\alpha\tau_0 = 1{\cdot}6/\pi$.

The delay per section, τ_0, is determined by the required bandwidth, the value of α by the slope at zero frequency and the number of sections, n, by the requirement that $\operatorname{cosech}\frac{1}{4}n\pi\alpha\tau_0$ is much less than unity in order that the ripples shall be negligible. In practice there is little advantage in making $\frac{1}{4}n\pi\alpha\tau_0$ greater than about 6, because the ripples will then be sufficiently small compared with other non-uniformities.

The generation of a vestigial-sideband signal requires two amplifying chains of identical phase and amplitude characteristics. It is difficult to produce such amplifiers more uniform in gain than ±1·0% and in phase than ±1·0°. With these limits the relative attenuation of the 'suppressed' sideband will be some 40 dB; Fig. 11 shows the ratio of the two sideband amplitudes as a function of dissimilarities of amplitude and phase. In order to ensure more than 40 dB relative attenuation of the suppressed sideband it is desirable to incorporate a filter, but this presents no major problem since the power to be dissipated is relatively small and dispersion in the region of cut-off is unimportant, in view of the small amplitudes of the components to be suppressed.

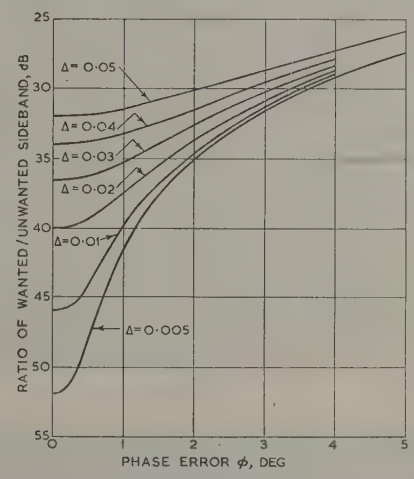


Fig. 11.—Ratio of sideband amplitude as a function of amplitude difference and phase difference between the two component double-sideband spectra.

Δ = Proportional error of amplitude.

The sections of delay line used in the quadrature network must have relatively uniform characteristics over the pass-band. The result of non-uniformity is twofold: there is the normal distortion which is due to the signal having passed through half the length of the delay line, and there are secondary effects which arise because the total signal comprises a number of separate contributions with differing delays and hence having differing amounts of distortion.

The effect of attenuation varying with frequency is to produce phase errors in the quadrature signal; on the other hand, varia-

tion of group delay will amplify errors in the signal spectrum, and both these effects will result in a failure to cancel completely the unwanted sideband components.

These effects are examined in detail in Sections 10.1 and 10.2 and it is concluded that, provided that the characteristic of the line is suitable for delay purposes, the additional distortions will not be serious.

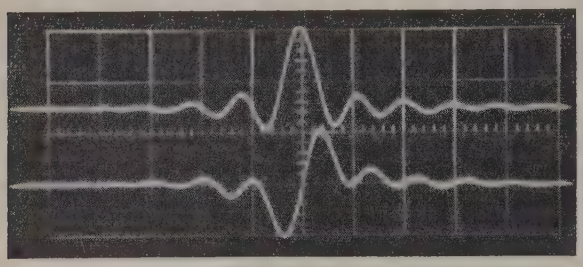
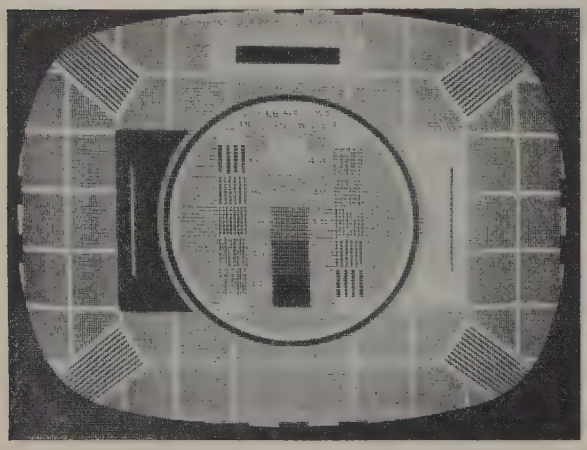
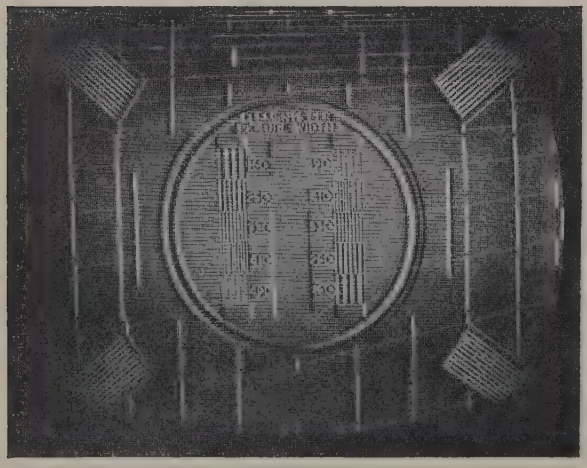


Fig. 12.—Quadrature signals obtained with a practical line.
Input signal (top) $\simeq \sin x/x$.



(*a*)



(*b*)

Fig. 13.—Outputs of a practical quadrature network.
(*a*) Output of the centre tap of a practical quadrature network with an input test card C.
(*b*) Quadrature output for the same input test card C.

Attenuation which is not a function of frequency will, c course, have no effect, provided that the correct signal level obtained at each tapping point.

Figs. 12 and 13(*a*) and (*b*) show results obtained with a prac tical quadrature network comprising eleven sections. Fig. 1 shows the output of the centre tap and the quadrature output fo an input waveform which is approximately $\sin x/x$. Similarly Fig. 13(*a*) shows a test card C from the centre tap and Fig. 13(*b* the corresponding quadrature signal. The picture monito brightness and contrast were increased for Fig. 13(*b*).

The two outputs of the quadrature unit were each used t modulate a 15 Mc/s carrier; the two carriers differed in phase b $\pi/2$ rad. The two modulated outputs were then added an supplied to a double-sideband receiver with an envelope detecto The resultant picture was imperceptibly different from Fig. 13(*a* and it was not considered worth while to include it in the pape

(6) THE TRANSMITTER

The method described above can be used to generate tw separate double-sideband signals which, when added, will pro duce a vestigial-sideband signal. In order to avoid the difficult of achieving efficient power amplification of a vestigial-sideban signal, two conventional transmitters using high-power modu lation may be used to generate each of the double-sideban signals, the addition taking place immediately before the aeria

One method worthy of consideration is the series connectio of the two transmitter output circuits and the aerial. Unde these conditions no power would be generated in the unwante sideband because of the phase opposition of the two transmitte at these frequencies. The transmitters would be presented wit an extremely high load impedance at these frequencies irrespe tive of the design load impedance presented at the othe frequencies.

Another method, perhaps more elegant, is to operate the tw transmitters in parallel by means of the hybrid network show in Fig. 14. By a suitable choice of dummy load and the charac teristic impedance of feeders, this network will present a constan impedance load at each transmitter, notwithstanding the fa that the wanted sideband and half the generated carrier powe are supplied to the aerial while the unwanted sideband powe

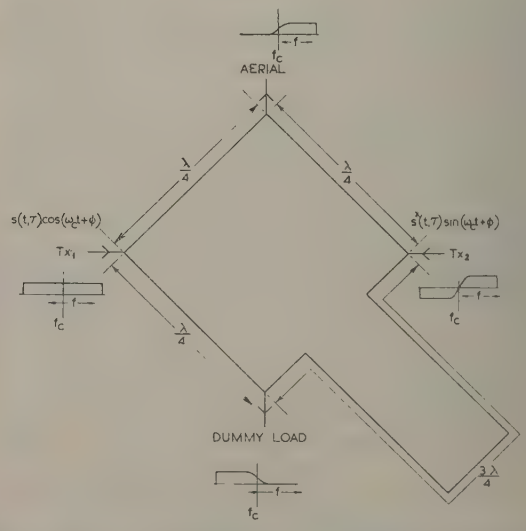


Fig. 14.—Hybrid network suitable for combining two double-sideban transmitters to produce a single-sideband signal.

and the other half of the generated carrier power are dissipated in the dummy load. In fact, the sum of the two spectra is supplied to the aerial and the difference of the two spectra is fed to the dummy load. In the case of a television system, one of the transmitters (Tx_1 in Fig. 14) is a normal double-sideband transmitter, but the second (Tx_2 in Fig. 14) will need to generate only the sidebands with no carrier and relatively little power in the sidebands close to the carrier. Thus the power of the second transmitter can be relatively small, although the total power output will be double that which would have been obtained from the transmitter T_1 had the unwanted sideband been removed by filtering.

Both these methods present the problem of maintaining sufficiently similar phase/frequency and amplitude/frequency characteristics in the two transmitters. A practical compromise might be achieved if the two transmitters have relative amplitude/frequency characteristics within $\pm 0{\cdot}25$ dB and relative phase/frequency characteristics within $\pm 3°$. This would permit the unwanted-sideband suppression to be some 30 dB. A simple filter in the aerial lead could be used to complete the suppression.

(7) OTHER APPLICATIONS

One advantage of the use of the quadrature network for obtaining vestigial-sideband signals is the complete freedom from restrictions in the choice of carrier frequency relative to the modulation frequencies. Because all the filtering is carried out at the modulation frequency, the carrier frequency can be changed without the need to switch vestigial-sideband filters.

A television signal having a bandwidth of 3 Mc/s could, if required, be modulated on to a carrier of, say, 0·4 Mc/s. Suitable design of the quadrature network can ensure that the resulting spectrum has zero amplitude at zero frequency, half amplitude at 0·4 Mc/s and full amplitude at 0·8 Mc/s extending to 3·4 Mc/s. This would permit the transmission of television signals through lines and amplifiers requiring little more than the video bandwidth but without the need for a d.c. component; wide-band transformers could thus be used without the need for clamp circuits and cathode-followers.

The method would also facilitate frequency changing for the purpose of multi-channel links.

(8) ACKNOWLEDGMENTS

The authors would like to express their thanks to the Director of Engineering of the British Broadcasting Corporation for permission to publish the paper, and to Mr. W. Proctor Wilson, the Head of the Research Department, Engineering Division, for many helpful discussions, for suggesting the alternative referred to in Section 4 and for providing the solutions to eqns. (10) and (11) in the paper.

Since writing the paper the authors have had their attention drawn to a work by M. Levy.[12] This dealt with the use of a delay line, connected as a 2-terminal impedance in a bridge circuit, for the production of a quadrature signal. Discontinuities were introduced into the line in order to provide reflections of appropriate amplitude, polarity and time delay at the input terminals.

(9) REFERENCES

(1) Honey, J. F., and Weaver, D. K.: 'Introduction to Single-Sideband Communications', *Proceedings of the Institute of Radio Engineers*, 1956, **44**, p. 1667.

(2) Norgaard, D. E.: 'The Phase-Shift Method of Single-Sideband Generation', *ibid.*, p. 1718.

(3) Weaver, D. K.: 'A Third Method of Generation and Detection of Single-Sideband Signals', *ibid.*, p. 1703.

(4) Dome, R. B.: 'Wide-Band Phase-Shift Networks', *Electronics*, December, 1946, **19**, p. 112.

(5) Luck, D. G. C.: 'Properties of Some Wide-band Phase-Splitting Networks', *Proceedings of the Institute of Radio Engineers*, 1949, **37**, p. 147.

(6) Saraga, W.: 'The Design of Wide-Band Phase-Splitting Networks', *ibid.*, 1950, **38**, p. 754.

(7) Weaver, D. K.: 'Design of *RC* Wide-Band 90° Phase-Difference Network', *ibid.*, 1954, **42**, p. 671.

(8) Gabor, D.: 'Theory of Communication', *Journal I.E.E.*, 1946, **93**, Part III, p. 429.

(9) Linke, J. M.: 'A Variable Time-Equalizer for Video-Frequency Waveform Correction', *Proceedings I.E.E.*, Paper No. 1252 R, April, 1952 (**99**, Part IIIA, p. 427).

(10) Nyquist, H., and Pfleger, K. W.: 'The Effect of the Quadrature Component in Single-Sideband Transmission', *Bell System Technical Journal*, 1940, **19**, p. 63.

(11) 'Single-Sideband Issue', *Proceedings of the Institute of Radio Engineers*, December, 1956, **44.**

(12) Levy, M.: 'The Impulse Response of Electrical Networks, with special reference to the Use of Artificial Lines in Network Design', *Journal I.E.E.*, 1943, **90**, Part III, p. 153.

(10) APPENDICES

(10.1) Effect of a Non-Uniform Delay/Frequency Characteristic of the Delay Line

Let the mean delay of each section of the line be τ_0 and allow the actual delay at any particular frequency to be $\tau_0(1 + k)$. If the input signal $s_1(t) = \cos(\omega t + \phi)$, the output, $s_2(t)$, will be

$$s_2(t) = A \sum_{x=-(n-1)/2}^{(n+1)/2} \cos \left\{ \omega \left[t - \frac{\tau_0(n + 2x - 1)(1 + k)}{2} \right] + \phi \right\} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x - 1)}{4} \quad . \quad (24)$$

This can be expanded as

$$s_2(t) = -\sin \left\{ \omega \left[t - \frac{n\tau_0(1 + k)}{2} \right] + \phi \right\} A \sum_{x=-(n-1)/2}^{(n+1)/2} \sin \frac{(1 + k)\omega\tau_0(2x - 1)}{2} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x - 1)}{4} + \cos \left\{ \omega \left[t - \frac{n\tau_0(1 + k)}{2} \right] + \phi \right\} A \sum_{x=-(n-1)/2}^{(n+1)/2} \cos \frac{(1 + k)\omega\tau_0(2x - 1)}{2} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x - 1)}{4} \quad . \quad . \quad . \quad . \quad (25)$$

The second term sums to zero, leaving only the first term.

The delayed input signal, $s_1(t, \tau)$, which is available at the centre tap of the line, can be expressed as

$$s_1(t, \tau) = \cos \left\{ \omega \left[t - \frac{n\tau_0(1 + k)}{2} \right] + \phi \right\} \quad . \quad . \quad (26)$$

Thus eqn. (25) can be expressed as

$$s_2(t) = s_1^*(t, \tau) A \sum_{x=-(n-1)/2}^{(n+1)/2} \sin \frac{\omega\tau_0(2x - 1)(1 + k)}{2} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x - 1)}{4} \quad . \quad . \quad . \quad . \quad (27)$$

Both $s_1(t, \tau)$ and $s_1^*(t, \tau)$ have delay distortions equal to $\frac{1}{2}kn\tau_0$, which is equal to the accumulated delay distortion of half the

number of sections in the line. This will result in the received and detected modulation having a delay distortion equal to $\frac{1}{2}kn\tau_0$; therefore the non-uniformity of each section must be kept sufficiently small for $\frac{1}{2}kn\tau_0$ to be acceptable. In addition to the above, the angle of the sine term in the series modifying the amplitude of $s_1^*(t, \tau)$ is increased to $\frac{1}{2}(1 + k)\omega\tau_0(2x - 1)$. This is equivalent to scaling the frequency response of the quadrature network by the ratio of $1/(1 + k)$. In practice, this effect will be negligible if the above-mentioned delay distortion of the transmitted signal is acceptable.

(10.2) The Effect of a Non-Uniform Amplitude/Frequency Characteristic of the Delay Line

Let the amplitude/frequency characteristic of each section of the delay line vary from the nominal or design value by a proportion Δ at any particular frequency. For an input $s_1(t) = \cos(\omega t + \phi)$ the output, $s_2(t)$, will be

$$s_2(t) = A \sum_{x=-(n-1)/2}^{(n+1)/2} \left[1 + \frac{\Delta(n + 2x - 1)}{2}\right] \cos\left\{\omega\left[t - \frac{(n + 2x - 1)\tau_0}{2}\right] + \phi\right\} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x - 1)}{4} \quad . \quad . \quad . \quad . \quad (28)$$

which on expansion becomes:

$$\begin{aligned} s_2(t) = &- A\left(1 + \frac{\Delta n}{2}\right) \sin\left[\omega\left(t - \frac{n\tau_0}{2}\right) + \phi\right] \sum_{x=-(n-1)/2}^{(n+1)/2} \sin \frac{\omega\tau_0(2x - 1)}{2} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x - 1)}{4} \\ &+ A\left(1 + \frac{\Delta n}{2}\right) \cos\left[\omega\left(t - \frac{n\tau_0}{2}\right) + \phi\right] \sum_{x=-(n-1)/2}^{(n+1)/2} \cos \frac{\omega\tau_0(2x - 1)}{2} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x - 1)}{4} \\ &- A\frac{\Delta}{2} \sin\left[\omega\left(t - \frac{n\tau_0}{2}\right) + \phi\right] \sum_{x=-(n-1)/2}^{(n+1)/2} (2x - 1) \sin \frac{\omega\tau_0(2x - 1)}{2} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x - 1)}{4} \\ &+ A\frac{\Delta}{2} \cos\left[\omega\left(t - \frac{n\tau_0}{2}\right) + \phi\right] \sum_{x=-(n-1)/2}^{(n+1)/2} (2x - 1) \cos \frac{\omega\tau_0(2x - 1)}{2} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x - 1)}{4} \quad . \quad . \quad . \quad . \quad (29) \end{aligned}$$

The second and third terms sum to zero, leaving the first and fourth terms.

The delayed input signal, $s_1(t, \tau)$, available at the centre tap of the line can be expressed as

$$s_1(t, \tau) = \left(1 + \frac{\Delta n}{2}\right) \cos\left[\omega\left(t - \frac{n\tau_0}{2}\right) + \phi\right] \quad . \quad (30)$$

Thus eqn. (29) can be rewritten as

$$\begin{aligned} s_2(t) = s_1^*(t, \tau) A \sum_{x=-(n-1)/2}^{(n+1)/2} \sin \frac{\omega\tau_0(2x - 1)}{2} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x - 1)}{4} \\ + s_1(t, \tau)\left(\frac{\Delta}{2 + \Delta n}\right) \times A \sum_{x=-(n-1)/2}^{(n+1)/2} (2x - 1) \cos \frac{\omega\tau_0(2x - 1)}{2} \operatorname{cosech} \frac{\pi\alpha\tau_0(2x - 1)}{4} \quad . \quad . \quad . \quad . \quad (31) \end{aligned}$$

Both the delayed input signal, $s_1(t, \tau)$, and its quadrature, $s_1^*(t, \tau)$, have a non-uniformity of amplitude equal to $\frac{1}{2}\Delta n$ resulting from the accumulated effect of half the number of sections in the line. Thus the non-uniformity of the amplitude/frequency characteristic of each section of the line should be small enough to ensure that the signal is not appreciably distorted after having travelled through half the length of the line.

A secondary effect of the non-uniform amplitude/frequency characteristic is to produce an output comprising the required quadrature component and a distortion component proportional in amplitude to the non-uniformity, Δ. This component is plotted in Fig. 15 in terms of Δ as a function of frequency for

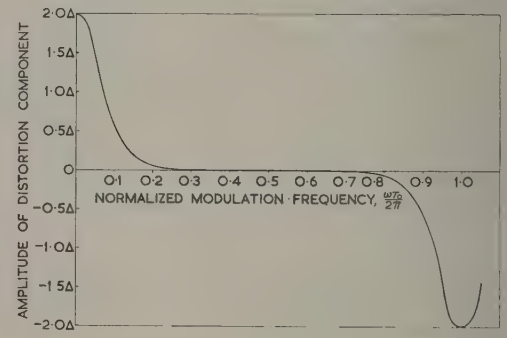


Fig. 15.—Distortion component caused by a non-uniform amplitude/frequency characteristic of line plotted as a function of frequency.
Δ = Proportional non-uniformity.

a line consisting of 15 sections with $\pi\alpha\tau_0 = 1{\cdot}6$ and $A = \sinh \frac{1}{4}\pi\alpha\tau_0$. These are the same conditions for the computed frequency response of the undistorted quadrature network plotted in Fig. 10(*c*). The distortion component has a maximum value of 2Δ at modulation frequencies around zero, rapidly reduces to negligible amplitude and rises to a negative peak of -2Δ at a normalized frequency $\omega\tau_0/2\pi = 1{\cdot}0$. Examination of Fig. 10(*c*) shows that this quadrature network is suitable for use only up to a maximum normalized modulation frequency of approximately 0·75; therefore we need consider only the distortion component for values of $\omega\tau_0/2\pi$ less than 0·75.

The effect of this distortion component in vestigial-sideband transmission is to produce a double-sideband spectrum in quadrature with the vestigial sidebands and having an amplitude 2Δ relative to the vestigial-sideband components near the carrier frequency. In practice, Δ will be less than 0·5% in order that the accumulated effect of half the number of sections in the line does not appreciably distort the delayed modulation signal $s_1(t, \tau)$.

[The discussion on the above paper will be found on page 281.]

.1.376.223

The Institution of Electrical Engineers
Paper No. 3051 E
Jan. 1960

RECTIFIER MODULATORS WITH FREQUENCY-SELECTIVE TERMINATIONS

With Particular Reference to the Effect of Even-Order Modulation Products

By D. P. HOWSON, M.Sc., Graduate, and Professor D. G. TUCKER, D.Sc., Member.

The paper was first received 8th October, 1958, and in revised form 28th April, 1959. It was published in January, 1960, and was read before the ELECTRONICS AND COMMUNICATIONS SECTION *11th January, 1960.)*

SUMMARY

Solutions are obtained for the modulation-product currents in ries- and shunt-type rectifier modulators by separating out an quation for each possible modulation-product frequency and by ssuming (*a*) square-wave switching of non-reactive rectifiers, and) signal-path terminating impedances which are a pure but not ecessarily constant resistance—or zero or infinite—at each modu- tion-product frequency involved. The conditions under which here are no even-order modulation-product currents are determined, nd the influence of such currents when they do exist is examined. xperimental results confirm the theory. It is shown that all the orking and conclusions apply exactly to the ring modulator, pro- ided that terminations required for odd-order modulation-product requencies are applied in the output loop and those required for ven-order frequencies are applied in the input.

(1) SERIES AND SHUNT TYPE MODULATORS

(1.1) Introduction

The calculation of the modulation products in a rectifier nodulator of the series or shunt single-balanced types shown in igs. 1 and 2 can be performed by writing out the circuit quation*

$$V = i[Z + r(t)] \quad . \quad . \quad . \quad . \quad . \quad (1)$$

lerived from the resultant common circuit shown in Fig. 3) eparately for every possible modulation product, $s\omega_p \pm \omega_q$, here s is any integer, and then solving the set of simultaneous quations.† This process has been discussed by a number of uthors, including Peterson and Hussey,[1] Peterson and lewellyn,[2] Belevitch,[3] and Tucker.[4] Except in some special ases,[5, 6] it has to be assumed (as in the present paper) that the ectifier impedance is a pure resistance, time-varying under the ction of a non-reactive carrier circuit, as represented by $r(t)$. is also assumed that the system is linear, i.e. the rectifier esistance function $r(t)$ is quite independent of the signal. If he signal level is low compared with the carrier level, this ssumption is reasonably justified; what happens when the ssumption is not made is discussed (for non-reactive circuits nly) by Belevitch[7] and by Tucker.[8] The time-varying resistance $t)$ is expanded as a Fourier series

$$r(t) = r_0 + \sum_{s=1}^{\infty} r_s \cos s\omega_p t \quad . \quad . \quad . \quad . \quad (2)$$

n the paper the assumption will be made that $r(t)$ is a square- ave function, so that s is always odd, and therefore, writing for s,

$$r_m = (-1)^{(m-1)/2} r_1/m \quad . \quad . \quad . \quad . \quad (3)$$

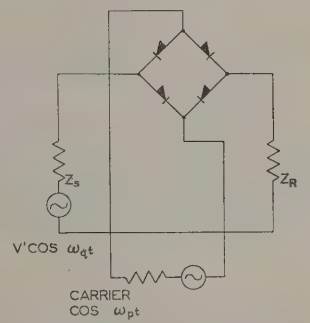


Fig. 1A.—Series modulator.

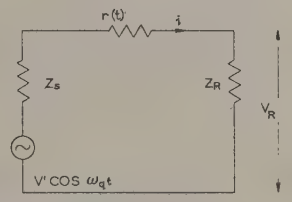


Fig. 1B.—Equivalent transmission circuit of series modulator.

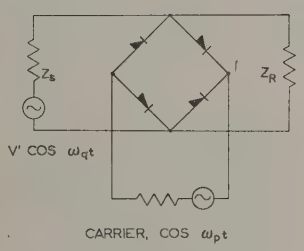


Fig. 2A.—Shunt modulator.

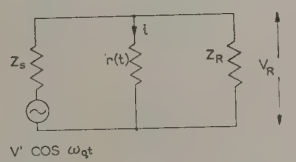


Fig. 2B.—Equivalent transmission circuit of shunt modulator.

It will also be assumed that Z is a pure resistance at all frequencies of modulation products where it is not zero or infinite, and it will therefore be replaced by the symbol R; as explained in Appendix 6.1 this leads to a considerable simplification of the equations and of their solution. At first glance the assumption

* A list of the symbols used in this paper and its companion will be found on ge 273.

† Note that the current-source equation, $I = v[Y + g(t)]$, could equally well be ed, in general, and would lead to a completely dual set of equations throughout. here are practical conditions of circuit termination where this alternative set of uations leads to simpler solutions, e.g. Case 6 in Section 1.4.

Mr. Howson is in the Department of Electrical Engineering, and Professor Tucker Professor of Electrical Engineering, University of Birmingham.

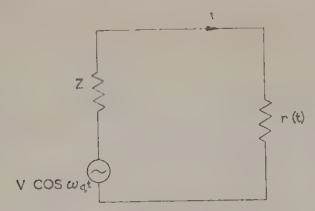


Fig. 3.—Resultant circuit for series or shunt modulator.

For series modulator:

$$V = V', Z = Z_S + Z_R, V_R = iZ_R.$$

For shunt modulator:

$$V = V'\frac{Z_R}{Z_S + Z_R}, Z = \frac{Z_S Z_R}{Z_S + Z_R} \quad V_R = iZ$$

may appear to involve unrealistic approximations, but consideration of Table 3 and Fig. 8, and of the experimental work described, will quickly show that many important practical cases can approximate very closely to these conditions.

If we expand eqn. (1) for each frequency, as set out in Appendix 6.1, we obtain the following results:

At frequency ω_q,

$$V = i_0(R_0 + r_0) + \tfrac{1}{2}\sum_{m=1,3,\ldots}(i_{m+} + i_{m-})r_m \quad . \quad . \quad (4)$$

At frequency $n\omega_p \pm \omega_q$, n odd,

$$0 = i_{n\pm}(R_{n\pm} + r_0) + \tfrac{1}{2}i_0 r_n + \tfrac{1}{2}\sum_{k=2,4,\ldots}(i_{k\pm}r_{|k-n|} + i_{k\mp}r_{|k+n|}) \quad (5)$$

At frequency $l\omega_p \pm \omega_q$, l even,

$$0 = i_{l\pm}(R_{l\pm} + r_0) + \tfrac{1}{2}\sum_{m=1,3,\ldots}(i_{m\pm}r_{|m-l|} + i_{m\mp}r_{|m+l|}) \quad (6)$$

The subscript notation is that used in previous literature, namely that $i_{n\pm}$ and $R_{n\pm}$ are the current and resistance at the frequency $n\omega_p \pm \omega_q$; i_0 and R_0 are the values at frequency ω_q. Both n and m are used for odd integers, but n is used to denote a particular integer, m being used in a more general sense, usually in summation formulae. Similarly l and k are used to denote even integers. It will be seen later, in eqns. (25)–(28), that all four symbols are necessary.

It is thus clear that even-order modulation products (i.e. those where the multiple of ω_p is even) can be produced in the modulator, even though r(t) contains only odd-order harmonics of ω_p. It is important to determine under what conditions even-order products are actually produced, what their effect is on the modulator performance, and under what conditions the calculation of modulation-product currents, and of conversion loss, can be made simple. There appears to be no previous publication of work on this subject of even-order products. The authors are, however, grateful to the Head of the Dr. Neher Laboratory of the Netherlands Postal and Telecommunications Service for permission to refer to an unpublished report on the subject by J. van der Graaf of that organization. This report reaches the same general conclusions as those of the present paper.

(1.2) Conditions for the Absence of Even-Order Modulation Products

To find the even-order currents as given in eqn (6), we must know the amplitudes of all the odd-order currents $i_{m\pm}$. These are given by eqn. (5) for each value n of m.

Thus we can rewrite eqn. (6) as

$$i_{l\pm} = \frac{1}{4(R_{l\pm} + r_0)}\left\{i_0 \sum_{m=1,3,\ldots}\left(\frac{r_m r_{|m-l|}}{R_{m\pm} + r_0} + \frac{r_m r_{|m+l|}}{R_{m\mp} + r_0}\right)\right.$$
$$+ \sum_{m=1,3,\ldots}\sum_{k=2,4,\ldots}\left[\frac{i_{k\pm}(r_{|k-m|}r_{|l-m|}) + i_{k\mp}(r_{|k+m|}r_{|l-m|}}{R_{m\pm} + r_0}\right.$$
$$\left.\left. + \frac{i_{k\pm}(r_{|k+m|}r_{|l+m|}) + i_{k\mp}(r_{|k-m|}r_{|l+m|})}{R_{m\mp} + r_0}\right]\right\} \quad . \quad . \quad . \quad .$$

Now, it is shown in Appendix 6.2 that

$$\sum_{m=1,3,\ldots}(r_m r_{|m-l|} + r_m r_{|m+l|}) = 0 \quad . \quad . \quad .$$

for l even, $l \neq 0$,

$$\sum_{m=1,3,\ldots}(r_{|k-m|}r_{|l-m|} + r_{|k+m|}r_{|l+m|}) = 0 \quad . \quad .$$

for k and l even, but $k \neq l$,

and $$\sum_{m=1,3,\ldots}(r_{|k+m|}r_{|l-m|} + r_{|k-m|}r_{|l+m|}) = 0 \quad . \quad (1$$

for k and l even.

Moreover, the order of the summations in the second te of eqn. (7) may be reversed. We then see that if

$$R_{m\pm} + r_0 = R_{m\mp} + r_0 = \text{constant} \quad . \quad . \quad (1$$

i.e. if R is constant for all odd-order modulation-product f quencies, so that the eqns. (8)–(10) can be applied to eqn. (then all the right-hand side of eqn. (7) becomes zero except the term in $k = l$ remaining from eqn. (9). We are then l with

$$i_{l\pm} = \frac{1}{4(R_{l\pm} + r_0)}\sum_{m=1,3,\ldots}\frac{i_{l\pm}(r^2_{|l-m|} + r^2_{|l+m|})}{R + r_0} \quad (12$$

i.e. $$i_{l\pm}\left[1 - \frac{\sum_{m=1,3,\ldots}(r^2_{|l-m|} + r^2_{|l+m|})}{4(R_{l\pm} + r_0)(R + r_0)}\right] = 0 \quad . \quad (12$$

Now it can be shown, along the lines of Appendix 6.2.6, th $\sum_{m=1,3,\ldots}(r^2_{|l-m|} + r^2_{|l+m|}) = \frac{1}{4}\pi^2 r_1^2$. The equation can therefo be written

$$i_{l\pm}\left[\frac{R_{l\pm}R + r_0(R_{l\pm} + R) + r_0^2 - \frac{\pi^2}{16}r_1^2}{(R_{l\pm} + r_0)(R + r_0)}\right] = 0 \quad (1$$

Now, if the back and forward resistances of the rectifiers are and r_f respectively,

$$r_0 = \frac{1}{2}(r_b + r_f) \quad \text{and} \quad r_1 = \frac{2}{\pi}(r_b - r_f)$$

so that $$r_0^2 - \frac{1}{16}\pi^2 r_1^2 = r_b r_f$$

Every term in the bracket multiplying $i_{l\pm}$ is therefore positi so that the value of the bracket cannot be zero. The o possible solution is therefore

$$i_{l\pm} = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad .$$

This means that, so long as R is a constant for all odd-or frequencies, no even-order currents flow. (Of course, to ensu that finite odd-order currents exist, it is also necessary to spec that R is not infinite at odd-order frequencies.)

It can easily be shown that this is a necessary, as well

ufficient, condition for the absence of even-order products. Thus, assume no even-order currents flow. Then, from eqn. (5),

$$i_{n\pm} = \frac{-\frac{1}{2}i_0 r_n}{R_{n\pm} + r_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

nd eqn. (6) becomes

$$\tfrac{1}{2}\sum_{m=1,3,\ldots}(i_{m\pm}r_{|m-l|} + i_{m\mp}r_{|m+l|}) = 0 \quad . \quad . \quad (15)$$

rovided that $R_{l\pm}$ is not infinite at any l,

e. $$\sum_{m=1,3,\ldots}\left(\frac{r_m r_{|m-l|}}{R_{m\pm} + r_0} + \frac{r_m r_{|m+l|}}{R_{m\mp} + r_0}\right) = 0 \quad . \quad . \quad (16)$$

hich is true, from eqn. (8), only if $R_{m\pm} + r_0 = R_{m\mp} + r_0 =$ onstant at all m. This is thus a necessary condition for the bsence of even-order products, except, of course, when R is nfinite at all even-order frequencies.

It is also evident from the above working that the even-order urrents are very nearly zero if at all odd-order frequencies

$$R_{m\pm} \ll r_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

ince this condition satisfies the conditions required to make he various series sum to zero.

It is thus concluded that the conditions for the absence of ven-order modulation-product currents are

(*a*) The resultant terminating impedance of the modulator (i.e. Z in Fig. 3) is a constant resistance at all odd-order frequencies.
(*b*) The resultant terminating impedance is infinite at all even-order frequencies.
(*c*) The resultant terminating impedance is very small at all odd-order frequencies compared with the mean value of the rectifier resistance.

(1.3) Conversion-Loss Calculations

Provided that the conditions stated in the previous Section or the absence of even-order products are met, the current at he wanted sideband frequency (assumed to be $\omega_p + \omega_q$) is asily determined. Thus, for condition (*a*), we make $R_{n\pm}$ in qn. (14) equal to a constant, say R_{1+}, and then substitute qn. (14) into eqn. (4). Using the well-known summation

$$1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots = \frac{\pi^2}{8}$$

e obtain

$$i_0 = \frac{V(R_{1+} + r_0)}{(R_{1+} + r_0)(R_0 + r_0) - \pi^2 r_1^2/16}$$

n substituting this back into eqn. (14), we obtain the solution

$$i_{1+} = \frac{-\frac{1}{2}Vr_1}{(R_{1+} + r_0)(R_0 + r_0) - \pi^2 r_1^2/16} \quad . \quad . \quad (18)$$

he absence of even-order terms reduces the calculation to ese simple steps, and this is, indeed, the main importance of nowing when even-order terms are absent. It is seen, too, at the relative amplitudes of the various modulation products re, in this modulator,

$$i_{1\pm} : i_{3\pm} : i_{5\pm} : \text{etc.} = 1 : \tfrac{1}{3} : \tfrac{1}{5} : \text{etc.} \quad . \quad . \quad (19)$$

If the modulator is of series type, three ways of maintaining he constant resistance at odd-order frequencies may be considered; practical circuits which approximate closely to these ases are discussed in Appendix 6.4 and illustrated in Fig. 8.

Case 1.—R_S and R_R each maintained constant over the whole frequency range. Then it is shown in Appendix 6.3 that the conversion loss is a minimum when $R_S = R_R = \rho/2$, where $\rho = \sqrt{(r_f r_b)}$, r_f and r_b being respectively the forward and backward resistance of the rectifiers, and that the minimum conversion loss is a ratio $\pi(1 + 2/n)$, where $n = \sqrt{(r_b/r_f)}$, or, in decibels, $9{\cdot}92 + 17{\cdot}4/n$ dB.

Case 2.—R_S and R_R each maintained constant over the whole frequency range, except that R_R is zero at frequency ω_q. Eqn. (18), expressed in terms of r_f and r_b becomes, for this case,

$$i_{1+} = \frac{-V(r_b - r_f)/\pi}{R_S(R_S + R_R) + \frac{1}{2}(r_f + r_b)(2R_S + R_R) + r_f r_b} \quad (20)$$

and the minimum conversion loss, assuming $R_S = R_R = R$, occurs when $R = \rho/\sqrt{2}$, and it is a ratio $\frac{3}{4}\pi(1 + 1{\cdot}89/n)$, or $7{\cdot}46 + 16{\cdot}4/n$ dB.

Case 3.—$R_S = 0$ and $R_R = R$ at frequency $\omega_p + \omega_q$; $R_R = 0$ and $R_S = R$ at frequency ω_q; $R_S + R_R =$ constant (R) at all odd-order frequencies. It is shown in Appendix 6.3 that the minimum conversion loss occurs when this constant resistance $R = \rho$, and it is a ratio $\frac{1}{2}\pi(1 + 2/n)$, or $3{\cdot}92 + 17{\cdot}4/n$ dB.

If the modulator is of shunt type, there are three corresponding ways:

Case 1.—R_S and R_R each maintained constant. The conversion loss is a minimum when $R_S = R_R = 2\rho$, and the minimum conversion loss is a ratio $\pi(1 + 2/n)$ as in the series case.

Case 2.—R_S and R_R each maintained constant over the whole frequency range, except that R_R is infinite at frequency ω_q. The minimum conversion loss occurs, assuming $R_S = R_R = R$, when $R = \rho\sqrt{2}$, and it is a ratio $\frac{3}{4}\pi(1 + 1{\cdot}89/n)$ as for the corresponding series modulator. (Note that this case is really complementary rather than analogous to Case 2 of the series modulator, since it is R_{1+} in eqn. (18) which is halved, and not R_0.)

Case 3.—$R_S = \infty$ and $R_R = R$ at frequency $\omega_p + \omega_q$, $R_R = \infty$ and $R_S = R$ at frequency ω_q; $R_S R_R/(R_S + R_R) =$ constant (R) at all odd-order frequencies. The minimum conversion loss occurs when $R = \rho$, and is a ratio $\frac{1}{2}\pi(1 + 2/n)$ as in the corresponding series case.

Calculations for condition (*b*) of Section 1.2 are easily made, but, since the terminating impedance can have resistive values of any magnitude at all odd-order frequencies, no simple general formula can be stated.

Case 4 is an important case which could occur in practice.

Series modulator with $R_S = R_{S0}$ at frequency ω_q, but zero at frequency $\omega_p + \omega_q$ and infinite at all other frequencies, and with $R_R = R_{R1+}$ at frequency $\omega_p + \omega_q$, but zero at frequency ω_q and infinite at all other frequencies.

Alternatively, shunt modulator with $R_S = R_{S0}$ at frequency ω_q, but infinite at all other frequencies; and with $R_R = R_{R1+}$ at frequency $\omega_p + \omega_q$, but infinite at all other frequencies.

In this case, no even-order currents flow, nor, in fact, can any currents exist except at ω_q and $\omega_p + \omega_q$. The very simple equations give, as shown in Appendix 6.3,

$$i_{1+} = \frac{-\frac{1}{2}Vr_1}{(R_{S0} + r_0)(R_{R1+} + r_0) - \frac{1}{4}r_1^2} \quad . \quad . \quad (21)$$

and the minimum conversion loss occurs when $R_{S0} = R_{R1+} = 0{\cdot}39\rho n$, and it has a ratio $2{\cdot}8(1 + 2{\cdot}6/n^2)$, or $8{\cdot}9 + 22{\cdot}6/n^2$ dB.

For condition (*c*), some important results, corresponding to common practical circuits, may be obtained.

Case 5.—Series modulator with $R_S =$ constant at all frequencies, $R_R = R_{R1+}$ at frequency $\omega_p + \omega_q$, but $= 0$ at all other frequencies. (This corresponds approximately to a shunt resonant circuit tuned to frequency $\omega_p + \omega_q$.) It is assumed that $R_S + R_R \ll r_0$, so that even-order terms may be neglected.

It is shown in Appendix 6.3 that

$$i_{1+} = \frac{-\frac{1}{2}Vr_1}{[R_S+R_{R1+}+r_0]\left[R_S+r_0-\frac{\frac{1}{4}r_1^2(\frac{1}{4}\pi^2-1)}{R_S+r_0}\right]-\frac{1}{4}r_1^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (22a)$$

$$= \frac{-V(r_b-r_f)/\pi}{\left\{R_S(R_S+R_{R1+})+\frac{1}{2}(r_f+r_b)(2R_S+R_{R1+})+\frac{1}{4}(r_f+r_b)^2-\frac{(r_b-r_f)^2}{\pi^2}-\frac{[R_S+R_{R1+}+\frac{1}{2}(r_f+r_b)]}{[R_S+\frac{1}{2}(r_f+r_b)]}\frac{\frac{1}{4}\pi^2-1}{\pi^2}(r_b-r_f)^2\right\}} \quad (22b)$$

But we are given that $R_S + R_{R1+} \ll \frac{1}{2}(r_f + r_b)$, so that

$$\frac{(r_b-r_f)^2}{\pi^2}\left[1+(\tfrac{1}{4}\pi^2-1)\frac{R_S+R_{R1+}+\frac{1}{2}(r_f+r_b)}{R_S+\frac{1}{2}(r_f+r_b)}\right] \simeq \tfrac{1}{4}(r_b-r_f)^2$$

and so

$$i_{1+} \simeq \frac{-V(r_b-r_f)/\pi}{R_S(R_S+R_{R1+})+\frac{1}{2}(r_f+r_b)(2R_S+R_{R1+})+r_fr_b} \quad (23)$$

This is exactly the same expression as for Case 2 of condition (*a*) of the series modulator [eqn. (20)], which is a different circuit having R_R constant except for a zero value at frequency ω_q only. But the present case requires the restriction $R_S + R_R \ll r_0$, which is not necessary for the former case.

The magnitude of the errors involved in assuming the absence of even-order products in this case is examined in Section 1.4.

(1.4) Calculations when Even-Order Products are Present

We will now proceed to examine what happens when the even-order modulation-product currents are not caused to be absent.

From eqn. (6),

$$i_{l\pm} = \frac{-1}{2(R_{l\pm}+r_0)}\sum_{m=1,3,\ldots}(i_{m\pm}r_{|m-l|}+i_{m\mp}r_{|m+l|}) \quad (24)$$

Therefore, if $R_{l+} = R_{l-}$,

$$i_{l+}r_{|l-n|}+i_{l-}r_{|l+n|} = \frac{-1}{2(R_l+r_0)}\sum_{m=1,3,\ldots}[i_{m+}(r_{|l-m|}r_{l-n|}+r_{|l+m|}r_{|l+n|}) + i_{m-}(r_{|l-m|}r_{|l+n|}+r_{|l+m|}r_{|l-n|})] \quad (25)$$

This can now be applied to the right-hand term in eqn. (5). Note that it is shown in Appendix 6.2 that

$$\sum_{k=2,4,\ldots}(r_{|k-m|}r_{|k-n|}+r_{|k+m|}r_{|k+n|}) = -r_mr_n \quad . \quad (26)$$

m and n odd, $m \neq n$

and that

$$\sum_{k=2,4,\ldots}(r_{|k-m|}r_{|k+n|}+r_{|k+m|}r_{|k-n|}) = -r_mr_n \quad . \quad (27)$$

m and n odd.

Therefore, provided that R_l is constant at all even-order frequencies, so that $1/2(R_l+r_0)$ can be taken outside the summation, the right-hand term of eqn. (5) is, for frequency $n\omega_p + \omega_q$,

$$\tfrac{1}{2}\sum_{k=2,4,\ldots}(i_{k+}r_{|k-n|}+i_{k-}r_{|k+n|}) = \frac{-1}{4(R_l+r_0)}\left[-\sum_{\substack{m=1,3,\ldots\\ m\neq n}}i_{m+}r_mr_n+i_{n+}\sum_{k-2,4,\ldots}(r_{|k-n|}^2+r_{|k+n|}^2)-\sum_{m=1,3,\ldots}i_{m-}r_mr_n\right] \quad (28)$$

The middle term of the right-hand side derives from the remainder of eqn. (26) when $m = n$.

It is also shown in Appendix 6.2 that

$$\sum_{k=2,4,\ldots}(r_{|k-n|}^2+r_{|k+n|}^2) = \tfrac{1}{4}\pi^2r_1^2-r_n^2 \quad . \quad . \quad (29)$$

The term $i_{n+}r_n^2$ can now be included in the first summation, so that $m \neq n$ is no longer a restriction on it. Then, applying eqn. (4), eqn. (28) becomes

$$\frac{-1}{4(R_l+r_0)}[2i_0r_n(R_0+r_0)-2Vr_n+i_{n+}\tfrac{1}{4}\pi^2r_1^2] \quad (30)$$

Therefore eqn. (5) now gives, assuming $R_0 = R_l$,

$$i_{n\pm} = \frac{-\frac{1}{2}Vr_n}{(R_{n\pm}+r_0)(R_l+r_0)-\pi^2r_1^2/16} \quad . \quad . \quad (31)$$

This is the general solution of the modulator for odd-order products, provided that the terminating impedance is a constant resistance at all even-order frequencies including ω_q. It will be observed that it is the same form as that in eqn. (18) for the case where the terminating impedance is a constant resistance at all odd-order frequencies, so that no even-order currents are produced.*

This general solution for constant resistance at even-order frequencies is very useful in at least one important practical case. In Section 1.3 [eqns. (22) and (23)] the following arrangement was discussed:

Case 5.—Series modulator with R_S = constant at all frequencies; $R_R = R_{R1+}$ at frequency $\omega_p + \omega_q$ (assumed to be the wanted sideband) but zero at all other frequencies. By the previous working it was necessary to assume $R_S + R_R \ll r_0$ in order to neglect even-order terms. But, by the working above we see that in this arrangement the terminating resistance is constant at all even-order frequencies (including ω_q), and from eqn. (31) the wanted sideband current is

$$i_{1+} = \frac{-\frac{1}{2}Vr_1}{(R_S+R_{R1+}+r_0)(R_S+r_0)-\pi^2r_1^2/16} \quad . \quad . \quad (31)$$

$$= \frac{-V(r_b-r_f)/\pi}{R_S(R_S+R_{R1+})+\frac{1}{2}(r_f+r_b)(2R_S+R_{R1+})+r_br_f} \quad (31)$$

which is identical with the expression previously obtained as eqn. (23), but without the requirement that $R_S + R_R \ll r_0$.

It is interesting to examine what error is introduced into the calculation of this circuit by merely ignoring even-order currents, i.e. we shall compare values of conversion loss using the correct calculation [eqn. (33)] and the calculation ignoring even-order currents [eqn. (22)]. Take $n = \sqrt{(r_b/r_f)} = 100$. The results are given in Table 1 for various values of $R_S = R_{R1+} = R$.

The error caused by the neglect of even-order currents is thus quite considerable, even when R is small.

Case 6.—Series modulator with $R_S = R_{S0}$ at frequency ω

* The results of eqns. (18) and (31) may be combined in the following very useful form. The solution of the circuit for the current at wanted sideband frequency, $\omega_p + \omega_q$, is

$$i_{1+} = \frac{-\frac{1}{2}Vr_1}{(R_{1+}+r_0)(R_0+r_0)-\pi^2r_1^2/16}$$

provided that either $R = R_{1+}$ at all odd-order frequencies, or $R = R_0$ at all even-order frequencies.

In the first condition, the value of R at even-order frequencies other than ω_q is immaterial, as there are no currents at these frequencies. In the second condition the value of R at odd-order frequencies other than $\omega_p + \omega_q$ is immaterial, as the odd-order currents do not interact.

Table 1

RELATIONSHIP BETWEEN RESISTANCE AND CONVERSION LOSS

R	Conversion loss from eqn. (33) (i.e. correct)	Conversion loss from eqn. (22) (i.e. approximate)
	dB	dB
2ρ	7·7	5·9
$\rho/\sqrt{2}$	7·6 (minimum condition)	5·7
$\rho/10$	8·0	6·2
$\rho/100$	11·9	10·7

but zero at all other frequencies; and with $R_R = R_{R1+}$ at frequency $\omega_p + \omega_q$ but zero at all other frequencies. In this case, none of the previous rigorous solutions can be applied, although if $R_S + R_R \ll r_0$, eqn. (23) is obtained as an approximate solution. The circuit can be solved by a laborious working out of the complete set of eqns. (4), (5) and (6), and then a complete answer is obtained, including the magnitudes of all even-order currents.[9] But such a complete solution is rarely required; since the terminations are of zero impedance at all frequencies other than ω_q and $\omega_p + \omega_q$, no voltage can be produced at any other frequency, so the magnitude of the currents is of little interest. In these circumstances a solution may be quickly obtained for i_{1+} by using the equation

$$i = V_1 g(t) \qquad . \quad . \quad . \quad . \quad . \quad (34)$$

where V_1 is the voltage across the rectifier $r(t)$, and $g(t) = 1/r(t)$. Noting that V_1 exists only for frequencies ω_q and $\omega_p + \omega_q$, the two simple equations for these frequencies are easily solved, giving

$$i_{1+} = \frac{-\frac{1}{2}Vg_1}{(R_{S0}g_0 + 1)(R_{R1+}g_0 + 1) - \frac{1}{4}g_1^2 R_{S0}R_{R1+}} \qquad (35)$$

The minimum conversion loss occurs when $R_{S0} = R_{R1+} = 2{\cdot}6\rho/n$, and is given by the ratio $2{\cdot}8\ (1 + 2{\cdot}6/n^2)$, or $8{\cdot}9 + 22{\cdot}6/n^2$ dB.

In this circuit, the error caused by neglecting even-order currents in the normal equations is very large. Table 2 shows the effect for $n^2 = 10^3$, $\rho^2 = 10^7$ (ohms).[2]

Table 2

VALUES OF CONVERSION LOSS

R	Correct conversion loss	Conversion loss ignoring even-order currents
ohms	dB	dB
1 000	12·2	1·0
250	8·9	4·7
50	13·7	12·6
10	24·8	24·5

When the even-order currents are neglected, the optimum conversion loss is calculated as $0{\cdot}87 + 22{\cdot}7/n$ dB for $R_{S0} = R_{R1+} = \rho$ which is a very different result from the correct one.

(2) THE RING MODULATOR

This rather more complicated circuit, shown in Fig. 4A, can be reduced to almost the same equations as the series and shunt modulators. Referring to the equivalent transmission circuit

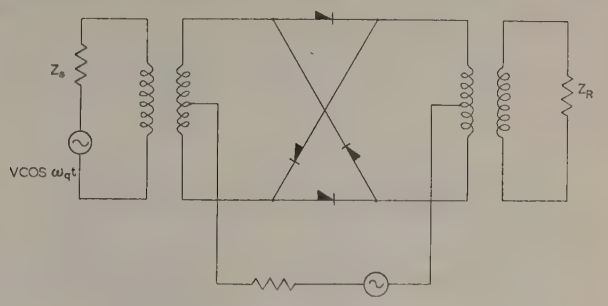


Fig. 4A.—The ring modulator.

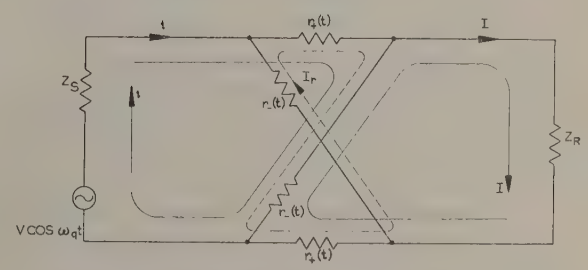


Fig. 4B.—Equivalent transmission circuit of ring modulator.

shown in Fig. 4B, where i is the current in the input loop Z_S, $r_+(t)$, and $r_-(t)$; I is the current in the output loop Z_R, $r_+(t)$, and $r_-(t)$; and I_r is the current in the rectifier ring $2r_+(t)$ and $2r_-(t)$; we have loop equations as follows:

$$V = i[Z_S + r_+(t) + r_-(t)] + I_r[r_+(t) + r_-(t)] - Ir_-(t). \quad (36)$$

$$0 = -ir_-(t) - I_r[r_+(t) + r_-(t)] + I[Z_R + r_+(t) + r_-(t)] \quad (37)$$

$$0 = i[r_+(t) + r_-(t)] + 2I_r[r_+(t) + r_-(t)] - I[r_+(t) + r_-(t)] \quad (38)$$

where $$r_+(t) = r_0 + \sum_{m=1,3,\ldots}^{\infty} r_m \cos m\omega_p t \quad . \quad . \quad . \quad (39a)$$

and $$r_-(t) = r_0 - \sum_{m=1,3,\ldots}^{\infty} r_m \cos m\omega_p t \quad . \quad . \quad . \quad (39b)$$

Thus $$r_+(t) + r_-(t) = 2r_0 \qquad . \quad . \quad . \quad . \quad . \quad (40)$$

The assumption will be made that Z_S and Z_R are pure resistances (or zero or infinite) at every modulation-product frequency.

It is now important to observe that, as shown in Appendix 6.5, and by Kruse,[11] the balanced-bridge arrangement prevents all even-order currents (including frequency ω_q) from flowing in Z_R and prevents all odd-order currents from flowing in Z_S. Thus

$$\left.\begin{aligned} i_{n\pm} &= 0 \\ I_{l\pm} &= 0 \\ I_0 &= 0 \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (41)$$

where, as before, n is odd and l is even.

Thus, from eqns. (36) and (38) at frequency ω_q,

$$V = i_0(R_{S0} + r_0) + \tfrac{1}{2}\sum_{m=1,3,\ldots} (I_{m+} + I_{m-})r_m \quad . \quad . \quad (42)$$

and from eqns. (37) and (38) at frequency $n\omega_p \pm \omega_q$,

$$0 = I_{n\pm}(R_{Rn\pm} + r_0) + \tfrac{1}{2}i_0 r_n + \tfrac{1}{2}\sum_{k=2,4,\ldots} (i_{k\pm}r_{|k-n|} + i_{k\mp}r_{|k+n|})$$

$$. \quad . \quad . \quad . \quad (43)$$

Table 3

MINIMUM CONVERSION LOSS UNDER VARYING TERMINATING CONDITIONS

Case No.	Terminating conditions			Minimum conversion loss (dB)
	Series	Shunt	Ring	
1	$R_S + R_R$ constant at all odd-order frequencies.	$R_S R_R/(R_S + R_R)$ constant at all odd-order frequencies.	R_R constant at all odd-order frequencies.	$3{\cdot}92 + 17{\cdot}4/n$ (+6 dB for series and shunt modulators).
2	$R_S + R_R$ constant at all odd-order frequencies, $R_{R0} = 0$.	$R_S R_R/(R_S + R_R)$ constant at all odd-order frequencies; $R_{R0} = \infty$ (complementary to series condition).	R_R constant at all odd-order frequencies, $R_{S0} = \frac{1}{2}R_R$ or $= 2R_R$.	$7{\cdot}46 + 16{\cdot}4/n$ (not applicable to ring modulator).
3	$R_{S1+} = 0$, $R_{S0} = R$, $R_{R0} = 0$, $R_{R1+} = R$; $R_S + R_R = R$ at all odd-order frequencies.	$R_{S1+} = \infty$, $R_{S0} = R$, $R_{R0} = \infty$, $R_{R1+} = R$; $R_S R_R/(R_S + R_R) = R$ at all odd-order frequencies.	R_R constant at all odd-order frequencies ($=R$), $R_{S0} = R$, but R_S any value at other frequencies.	$3{\cdot}92 + 17{\cdot}4/n$.
4	$R_{S0} = R$, $R_{S1+} = 0$, $R_{R0} = 0$, $R_{R1+} = R$, $R_S + R_R = \infty$ at all other frequencies.	$R_{S0} = R$, $R_S = \infty$ at all other frequencies, $R_{R1+} = R$, $R_R = \infty$ at all other frequencies.	$R_{S0} = R$, $R_S = \infty$ at all even-order frequencies. $R_{R1+} = R$, $R_R = \infty$ at all odd-order frequencies.	$8{\cdot}9 + 22{\cdot}6/n^2$.
5	$R_{S0} = R$; $R_{S1+} = R$, $R_{R1+} = R$. $R_S + R_R = R$ at all even-order frequencies.	$R_{S0} = R$; $R_{S1+} = R$. $R_{R1+} = R$; $R_S R_R/(R_S + R_R) = R$ at all even-order frequencies (complementary to series condition).	R_S constant at all even-order frequencies. $R_{R1+} = 2R_S$ or $\frac{1}{2}R_S$, but R_R any value at other frequencies.	$7{\cdot}46 + 16{\cdot}4/n$ (not applicable to ring modulator).
6	$R_{S0} = R$, $R_S = 0$ at all other frequencies. $R_{R1+} = R$, $R_R = 0$ at all other frequencies.	$R_{S0} = R$, $R_{S1+} = \infty$ $R_S R_R/(R_S + R_R) = 0$ at all other frequencies. $R_{R0} = \infty$, $R_{R1+} = R$.	$R_{S0} = R$, $R_S = 0$ at all other even-order frequencies. $R_{R1+} = R$, $R_R = 0$ at all other odd-order frequencies.	$8{\cdot}9 + 22{\cdot}6/n^2$.
7	$R_S + R_R$ constant at all even-order frequencies.	$R_S R_R/(R_S + R_R)$ constant at all even-order frequencies.	R_S constant at all even-order frequencies.	$3{\cdot}92 + 17{\cdot}4/n$ (+6 dB for shunt and series modulators).
8	$R_{S0} = R$, $R_{R1+} = R$, $R_S + R_R = 0$ at all other odd-order frequencies, but $R_S + R_R = \infty$ at all other even-order frequencies.	$R_{S0} = R$, $R_{S1+} = \infty$; $R_{R0} = \infty$, $R_{R1+} = R$; $R_S R_R/(R_S + R_R) = 0$ at all other odd-order frequencies, but $= \infty$ at all other even-order frequencies.	$R_{S0} = R$, $R_S = \infty$ at all other even-order frequencies. $R_{R1+} = R$, $R_R = 0$ at all other odd-order frequencies.	$0{\cdot}87 + 22{\cdot}7/n$.

N.B. Where no value is specified for R_S or R_R at a particular modulation-product frequency, the value is immaterial. The conversion-loss figures assume $R_{S0} = R_{R1+}$.

and from eqns. (36) and (38) at frequency $l\omega_p \pm \omega_q$,

$$0 = i_{l\pm}(R_{Sl\pm} + r_0) + \tfrac{1}{2} \sum_{m=1,3,\ldots} (I_{m\pm r|m-l|} + I_{m\mp r|m+l|}) \quad (44)$$

The relation between these three eqns. (42), (43) and (44) and those for the series or shunt modulator (4), (5) and (6) is easily seen to be very close. They are, in fact, identical except that, since even- and odd-order currents flow in separate loops in the ring modulator, but both types of current flow in the one and only loop in the simplified series/shunt modulator circuit of Fig. 3, it is necessary to use I for odd-order and i for even-order currents, and to distinguish between R_S and R_R, in the ring-modulator equations. The consequence of this relationship is that all the conclusions reached and formulae obtained for the series and shunt modulators apply equally to the ring modulator, provided that any terminating condition specified for the single-loop series/shunt modulator at odd-order frequencies is applied in the output loop of the ring modulator (i.e. to R_R), and that any terminating condition specified for even-order frequencies is applied in the input loop (i.e. to R_S). For example, for even-order currents to be completely absent from the ring modulator it is necessary to specify that

(*a*) R_R is a constant resistance at all odd-order frequencies.
or (*b*) R_S is infinite at all even-order frequencies.
or (*c*) R_R is very small at all odd-order frequencies compared with the mean value of the rectifier resistance.

These three conditions correspond to those listed for the series/shunt modulator at the end of Section 1.2.

The way in which particular cases of series/shunt and ring modulators correspond is shown in Table 3. It should be noted that, since in the series/shunt modulator equations R_S and R_R have been combined, then in some cases (namely those where R_S and R_R are both finite at the wanted sideband frequency) the wanted sideband current or voltage is shared between R_S and R_R. In the conversion-loss figures it is assumed that the wanted sideband is $\omega_p + \omega_q$ and that R_{S0} and R_{R1+} are equal, so that in these cases the conversion loss is 6 dB more for series and shunt modulators than for the ring modulator. Note that in the Table, $n = \sqrt{(r_b/r_f)}$.

It should also be observed that in each of Cases 2 and 5, the shunt and series conditions are complementary rather than equivalent. Since for both Cases 2 and 5 the ring modulator has not equal source and load terminations at the input and output frequencies respectively, the conversion-loss figure does not strictly apply to it. Cases 1–6 correspond to the cases dealt with under these numbers in the earlier Sections, but in the Table the terminations have been defined in their strict theoretical form and not in the practical form which was the basis of the earlier discussion and of the experimental measurements described below. A selection of simple practical circuits which approximate to Cases 1 to 8 is shown in Fig. 8.

(3) EXPERIMENTAL RESULTS

Measurements were made on a series modulator set up as shown in Fig. 5. A battery-driven oscillator was used for the carrier supply in order to eliminate spurious couplings between

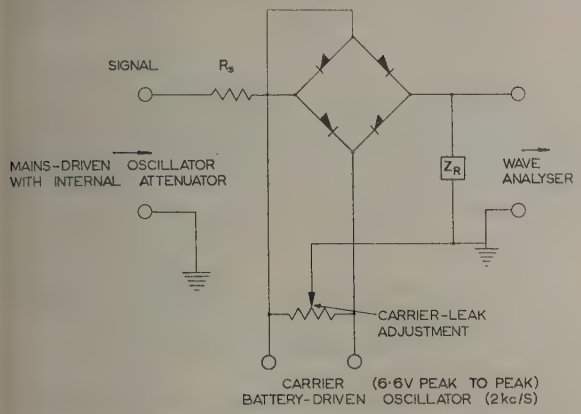


Fig. 5.—Experimental circuit.

the signal and carrier sources. The carrier oscillator had a square-wave output, but great care had to be taken to preserve equal 'mark' and 'space' durations, since inequality would have introduced even-order harmonics in the rectifier resistance function r(*t*); these would have produced even-order modulation products which would have obscured the even-order products caused by the terminating conditions.

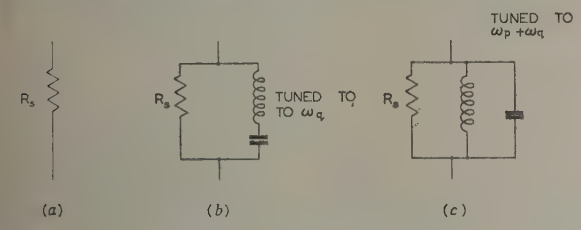


Fig. 6.—The three arrangements of Z_R in Fig. 5.

Three cases were studied (see Fig. 6), namely

(*a*) R_S = constant resistance at all frequencies and Z_R = constant resistance (= R_S) at all frequencies.

(*b*) R_S = constant resistance at all frequencies, and Z_R = series-resonant circuit, tuned to signal frequency (345 c/s in this case) and giving substantially zero impedance at this frequency, in parallel with a resistance (= R_S). The resonance was so sharp that the shunting effect of the resonant circuit at the frequencies of the various modulation products can be ignored (reactance at least $10R_S$).

(*c*) R_S = constant resistance at all frequencies, and Z_R = shunt resonant circuit, tuned to wanted sideband frequency ($\omega_p + \omega_q$), in parallel with a resistance (= R_S). This circuit was effectively a short-circuit at all the frequencies concerned, apart from $\omega_p + \omega_q$.

The theory of these three cases has been discussed in Sections 1.3 and 1.4. Case (*a*) corresponds to Section 1.3, condition (*a*), Case 1, and the minimum conversion loss of $9{\cdot}92 + 17{\cdot}4/n$ dB occurs when $R_S = R_R = \rho/2$. Case (*b*) corresponds to Section 1.3, condition (*a*), Case 2, and the conversion loss is calculated from eqn. (20). Case (*c*) corresponds to that dealt with on an approximate basis in Section 1.3, condition (*c*), and on an exact basis in Section 1.4, Case 5, eqn. (33); for the same value of R_S, the conversion loss is the same as in Case (*b*).

Now for the germanium rectifiers used, it was found that $r_f \simeq 75$ ohms, and $r_b \simeq 450$ kilohms for the carrier voltage used. Thus $n = 77{\cdot}5$ and $\rho = 5\,820$ ohms. The theoretical figures are compared with the experimental measurements of conversion loss, for the values of R_S indicated, in Table 4.

Table 4

THEORETICAL AND MEASURED CONVERSION LOSS

Case	R_S	Theoretical conversion loss	Measured conversion loss
	ohms	dB	dB
(*a*)	2900 (optimum)	10·14	10·1
(*b*)	1200	7·9	7·8
(*c*)	(i) 1100	7·9	7·9
	(ii) 2800	7·65	7·6

The agreement between theory and measurement is evidently within the limits of experimental accuracy.

In Case (*c*), theory shows that even-order currents should be produced. Fig. 7 shows the calculated values of odd- and even-

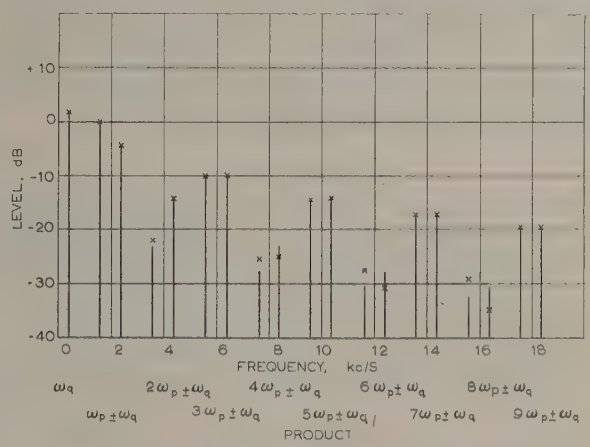


Fig. 7.—Graph of theoretical and measured currents for odd- and even-order modulation products in case (*c*).
Vertical lines are theoretical values; crosses indicate measured values.

order currents (for $R_S = 1\,100$ ohms) compared with the measured values. It will be seen that the agreement in odd-order currents is excellent; for even-order currents there are inevitably greater errors, owing to the difficulty of obtaining an exactly equal mark and space duration in the carrier wave, but the agreement is nevertheless sufficient to vindicate the theory.

In Case (*b*), theory shows that even-order currents should not exist. Experimental measurements confirmed this within the limits set by the residual even-order currents caused by unequal mark and space durations. The even-order currents were found to be less than 1% of the current at frequency $\omega_p + \omega_q$, and their amplitudes were all practically equal—as would be expected if their source was unequal mark and space durations, but not if they were produced by interactions as discussed in the paper.

(4) CONCLUSIONS

Throughout the work it has been assumed that r(*t*) is a square wave.

It has been shown that, in a series or shunt modulator (see Figs. 1, 2 and 3), even-order modulation-product currents do not exist if one of the following three conditions holds:

(*a*) Z is a constant resistance at the frequencies of all odd-order modulation products.
(*b*) Z is infinite at the frequencies of all even-order modulation products.
(*c*) Z is very small compared with the mean rectifier resistance at the frequencies of all odd-order products.

It has also been shown that a simple solution can be found for the odd-order currents if Z is a constant resistance at the frequencies of all even-order products (including ω_q) and is a pure resistance of any value, including zero and infinity, at the frequencies of all odd-order products. In this case, the current at any particular odd-order frequency where $Z = R_{n\pm}$ is the same as for case (*a*) above with the constant resistance at odd-order frequencies made equal to $R_{n\pm}$.

A number of examples of modulator arrangements where the theory enables the performance to be calculated have been described, and the effect of ignoring even-order currents where they exist has been shown to be often very large. Experimental measurements of three arrangements have confirmed the theory very closely.

It has been shown that the ring modulator corresponds exactly to the series/shunt modulator except that odd-order currents are confined to the output loop, and even-order currents to the input loop, and the terminating conditions discussed for the series/shunt modulator must therefore be applied to the appropriate end of the ring modulator.

(5) REFERENCES

(1) PETERSON, E., and HUSSEY, L. W.: 'Equivalent Modulator Circuits', *Bell System Technical Journal*, 1939, **18**, p. 32.

(2) PETERSON, L. C., and LLEWELLYN, F. B.: 'Performance and Measurement of Mixers in Terms of Linear Network Theory', *Proceedings of the Institute of Radio Engineers*, 1945, **33**, p. 458.

(3) BELEVITCH V.: 'Linear Theory of Bridge and Ring Modulator Circuits', *Electrical Communication*, 1948, **25**, p. 62.

(4) TUCKER, D. G.: 'Rectifier Modulators with Frequency-Selective Terminations', *Proceedings I.E.E.*, Paper No. 863 R, September, 1949 (**96**, Part III, p. 422).

(5) TUCKER, D. G.: 'Modulators, Frequency Changers and Detectors using Rectifiers with Frequency-Dependent Characteristics', *ibid.*, Paper No. 1139 R, September, 1957 (**98**, Part III, p. 394).

(6) BELEVITCH, V.: 'Theorié du Modulateur à Coupure Shunté par Capacité', *H.F.* (Brussels), 1952, **2**, p. 1.
'Effect of Rectifier Capacitance on the Conversion Loss of Ring Modulators', *Transactions of the Institute of Radio Engineers*, March, 1955, **CT-2**, **1**, p. 41.

(7) BELEVITCH, V.: 'Non-Linear Effects in Ring Modulators', *Wireless Engineer*, 1949, **26**, p. 177.
'Non-Linear Effects in Rectifier Modulators', *ibid.*, 1950, **27**, p. 130.
'Non-Linear Distortion in a Cowan Modulator', *ibid.*, p. 164.

(8) TUCKER, D. G.: 'Intermodulation Distortion in Rectifier Modulators', *ibid.*, 1954, **31**, p. 145.

(9) HOWSON, D. P.: Thesis, Electrical Engineering Department, University of Birmingham, 1958.

(10) KNOPP: 'Theory and Applications of Infinite Series' (Blackie, 1928), p. 234.

(11) KRUSE, S.: 'Theory of Rectifier Modulators', *Ericsson Technics*, 1939, **2**, p. 17.

(6) APPENDICES

(6.1) Expansion of Eqn. (1)

In eqn. (1), i is the sum of the currents at all modulation-product frequencies, i.e. at all frequencies $s\omega_p \pm \omega_q$, where s is any integer. The time-varying resistance $r(t)$ can be written as a Fourier series, eqn. (2), and therefore the product term $i.r(t)$ consists of the sum of products of two sinusoids,

$$i.r(t) = \left\{ \sum_{s=0}^{\infty} i_{s\pm} \cos \left[(s\omega_p \pm \omega_q)t + \theta_{s\pm} \right] \right\} \left(r_0 + \sum_{s=1}^{\infty} r_s \cos s\omega_p t \right) \quad (45)$$

In this notation, the subscript $s+$ means 'value at frequency $s\omega_p + \omega_q$'. It is clear that, in half the terms obtained from the product of two sinusoids, the phase angles are reversed in sign.

Therefore

$$\begin{aligned}
i.r(t) = {} & i_0 r_0 \cos(\omega_q t + \theta_0) + \tfrac{1}{2} i_{1+} r_1 \cos(\omega_q t + \theta_{1+}) + \tfrac{1}{2} i_{1-} r_1 \cos(\omega_q t - \theta_{1-}) \\
& + \tfrac{1}{2} i_{2+} r_2 \cos(\omega_q t + \theta_{2+}) + \tfrac{1}{2} i_{2-} r_2 \cos(\omega_q t - \theta_{2-}) + \dots \\
& + i_{1+} r_0 \cos[(\omega_p + \omega_q)t + \theta_{1+}] + \tfrac{1}{2} i_0 r_1 \cos[(\omega_p + \omega_q)t + \theta_0] + \tfrac{1}{2} i_{1-} r_2 \cos[(\omega_p + \omega_q)t - \theta_{1-}] \\
& + \tfrac{1}{2} i_{2+} r_1 \cos[(\omega_p + \omega_q)t + \theta_{2+}] + \tfrac{1}{2} i_{2-} r_3 \cos[(\omega_p + \omega_q)t - \theta_{2-}] + \dots \\
& + i_{1-} r_0 \cos[(\omega_p - \omega_q)t + \theta_{1-}] + \tfrac{1}{2} i_0 r_1 \cos[(\omega_p - \omega_q)t - \theta_0] + \tfrac{1}{2} i_{1+} r_2 \cos[(\omega_p - \omega_q)t - \theta_{1+}] \\
& + \tfrac{1}{2} i_{2+} r_3 \cos[(\omega_p - \omega_q)t - \theta_{2+}] + \tfrac{1}{2} i_{2-} r_1 \cos[(\omega_p - \omega_q)t + \theta_{2-}] + \dots \\
& + \dots \\
& \vdots \\
& + i_{s+} r_0 \cos[(s\omega_p + \omega_q)t + \theta_{s+}] + \tfrac{1}{2} i_0 r_s \cos[(s\omega_p + \omega_q)t + \theta_0] + \tfrac{1}{2} i_{1+} r_{s-1} \cos[(s\omega_p + \omega_q)t + \theta_{1+}] \\
& + \tfrac{1}{2} i_{1-} r_{s+1} \cos[(s\omega_p + \omega_q)t - \theta_{1-}] + \tfrac{1}{2} i_{2+} r_{s-2} \cos[(s\omega_p + \omega_q)t + \theta_{2+}] + \dots \\
& + \dots \\
& \vdots
\end{aligned} \quad (46)$$

Now, the other product term in eqn. (1) is iZ, and this is easily dealt with by expanding in this way:

$$iZ = \left\{\sum_{s=0}^{\infty} i_{s\pm}\cos\left[(s\omega_p \pm \omega_q)t + \theta_{s\pm}\right]\right\} Z$$
$$= \sum_{s=0}^{\infty} i_{s\pm}Z_{s\pm}\cos\left[(s\omega_p \pm \omega_q)t + \theta_{s\pm}\right] \quad . \quad . \quad (47)$$

where the value of Z at each frequency is put in separately.

Eqn. (1) is true at all instants of time, and therefore there must be a separate equilibrium condition for every frequency present. All terms of the same frequency are thus grouped together, and since $V = 0$ at all frequencies except ω_q, at all frequencies except ω_q the left-hand side of each equation is zero.

At frequency ω_q,

$$V\cos\omega_q t = i_0 Z_0\cos(\omega_q t + \theta_0) + i_0 r_0\cos(\omega_q t + \theta_0)$$
$$+ \tfrac{1}{2}i_{1+}r_1\cos(\omega_q t + \theta_{1+}) + \tfrac{1}{2}i_{1-}r_1\cos(\omega_q t - \theta_{1-})$$
$$+ \tfrac{1}{2}i_{2+}r_2\cos(\omega_q t + \theta_{2+}) + \ldots \quad . \quad . \quad . \quad (48)$$

At frequency $\omega_p + \omega_q$,

$$0 = i_{1+}Z_{1+}\cos\left[(\omega_p + \omega_q)t + \theta_{1+}\right] + i_{1+}r_0\cos\left[(\omega_p + \omega_q)t + \theta_{1+}\right] + \tfrac{1}{2}i_0 r_1\cos\left[(\omega_p + \omega_q)t + \theta_0\right] + \tfrac{1}{2}i_{1-}r_2\cos\left[(\omega_p + \omega_q)t - \theta_{1-}\right] + \tfrac{1}{2}i_{2+}r_1\cos\left[(\omega_p + \omega_q)t + \theta_{2+}\right] + \ldots \quad (49)$$

and, generally, at frequency $s\omega_p + \omega_q$,

$$0 = i_{s+}Z_{s+}\cos\left[(s\omega_p + \omega_q)t + \theta_{s+}\right] + i_{s+}r_0\cos\left[(s\omega_p + \omega_q)t + \theta_{s+}\right] + \tfrac{1}{2}i_0 r_s\cos\left[(s\omega_p + \omega_q)t + \theta_0\right] + \tfrac{1}{2}i_{1+}r_{s-1}\cos\left[(s\omega_p + \omega_q)t + \theta_{1+}\right] + \tfrac{1}{2}i_{1-}r_{s+1}\cos\left[(s\omega_p + \omega_q)t - \theta_{1-}\right] + \ldots \quad (50)$$

This system of equations may be solved for any specified and finite set of values of Z, but the working is usually laborious.

It is clear that a very great simplification results if it is specified that Z is a pure resistance at all frequencies of modulation products where it is not zero or infinite. This restriction means that all the phase angles ($\theta_{s\pm}$) become zero, so that the cosine terms may be dropped from the equations, leaving, for example, at frequency ω_q:

$$V = i_0 R_0 + i_0 r_0 + \tfrac{1}{2}i_{1+}r_1 + \tfrac{1}{2}i_{1-}r_1 + \tfrac{1}{2}i_{2+}r_2 + \ldots \quad (51)$$

If, further, it is assumed that r(t) is a symmetrical wave, so that $r_s = 0$ when s is even, then another simplification results, and the set of equations derived above become merely the set (4), (5) and (6), where the equilibrium for odd-order frequencies has been set out separately from that for even-order frequencies in order to show how the odd- and even-order terms interact.

(6.2) Proofs of Summations of Series

(6.2.1) **Case (a).**

To prove

$$\sum_{m=1,3,\ldots}^{\infty}(r_m r_{|m-l|} + r_m r_{|m+l|}) = 0 \quad . \quad . \quad (52)$$
for l even, $l \neq 0$.

Put $m = 2m' + 1$ and $l = 2l'$, so that m' and l' may take all integral values. Then

$$r_m = \frac{(-1)^{m'}}{2m'+1}r_1 \quad . \quad . \quad . \quad . \quad . \quad (53)$$

and the expression becomes

$$(-1)^{l'}r_1^2\sum_{m'=0}^{\infty}\left[\frac{1}{(2m'+1)(2m'-2l'+1)} + \frac{1}{(2m'+1)(2m'+2l'+1)}\right] \quad (54)$$

$$= (-1)^{l'}r_1^2\left[\sum_{m'=0}^{l'-1}\frac{1}{(2m'+1)(2m'-2l'+1)} + 2\sum_{m'=0}^{\infty}\frac{1}{(2m'+1)(2m'+2l'+1)}\right] \quad (55)$$

since

$$\sum_{m'=l'}^{\infty}\frac{1}{(2m'+1)(2m'-2l'+1)} = \sum_{m'=0}^{\infty}\frac{1}{(2m'+1)(2m'+2l'+1)} \quad . \quad . \quad . \quad . \quad (56)$$

Now

$$\sum_{m'=0}^{l'-1}\frac{1}{(2m'+1)(2m'-2l'+1)} = -\frac{1}{2l'}\sum_{m'=0}^{l'-1}\left(\frac{1}{2m'+1} + \frac{1}{2l'-2m'-1}\right)$$
$$= -\frac{1}{2l'}\left[\left(1 + \frac{1}{3} + \ldots + \frac{1}{2l'-1}\right) + \left(\frac{1}{2l'-1} + \frac{1}{2l'-3} + \ldots + \frac{1}{3} + 1\right)\right] \quad (57)$$

and

$$\sum_{m'=0}^{\infty}\frac{1}{(2m'+1)(2m'+2l'+1)} = \frac{1}{2l'}\left(1 + \frac{1}{3} + \ldots + \frac{1}{2l'-1}\right) \text{ (see Ref. 10)} \quad (58)$$

The two terms of eqn. (55) are therefore equal and of opposite sign, and the expression becomes zero.

(6.2.2) **Case (b).**

To prove

$$\sum_{m=1,3,\ldots}^{\infty}(r_{|k-m|}r_{|l-m|} + r_{|k+m|}r_{|l+m|}) = 0 \quad . \quad . \quad (59)$$
for k and l even, but $k \neq l$.

This may be rewritten as

$$\sum_{m=1,3,\ldots}^{\infty}\left[r_m r_{|m+(k-l)|} + r_m r_{|m-(k-l)|}\right] + \sum_{m=1-k}^{-1} r_{|m|}r_{|m+(k-l)|} - \sum_{m=1}^{k-1} r_m r_{|m-(k-l)|} \quad (60)$$

The first summation is zero, by eqn. (52).

The second and third summations expand to

$$(r_{|k-1|}r_{|l-1|} + \ldots + r_1 r_{|k-l-1|}) - (r_1 r_{|k-l-1|} + \ldots + r_{|k-1|}r_{|l-1|})$$

which is also zero.

(6.2.3) **Case (c).**

To prove

$$\sum_{m=1,3,\ldots}^{\infty}(r_{|k+m|}r_{|l-m|} + r_{|k-m|}r_{|l+m|}) = 0 \quad . \quad (61)$$
for k and l even,
but $k + l \neq 0$.

This may be rewritten as

$$\sum_{m=1,3,\ldots}^{\infty} [r_m r_{|m-(k+l)|} + r_{m'} r_{|m+(k+l)|}] + \sum_{m=1-k}^{-1} r_{|m|} r_{|m+(k+l)|} - \sum_{m=1}^{k-1} r_m r_{|m-(k+l)|} \quad (62)$$

The first summation is zero, by eqn. (52).

The second and third summations expand to

$$[r_{|k-1|} r_{|l+1|} + \ldots + r_1 r_{|1-(k+l)|}] - [r_1 r_{|1-(k+l)|} + \ldots + r_{|k-1|} r_{|l+1|}]$$

which is also zero.

(6.2.4) Case (d).

To prove

$$\sum_{k=2,4,\ldots}^{\infty} (r_{|k-m|} r_{|k-n|} + r_{|k+m|} r_{|k+n|}) + r_m r_n = 0 \quad (63)$$

for m and n odd, $m \neq n$.

The expression may be rewritten as

$$\sum_{M=1,3,\ldots}^{\infty} [r_M r_{|M-(n-m)|} + r_M r_{|M+(n-m)|}] + \sum_{M=2-m}^{-1} r_{|M|} r_{|M-(n-m)|} - \sum_{M=1}^{m-2} r_M r_{|M+(n-m)|} \quad (64)$$

where the term $r_m r_n$ has been included in $r_M r_{|M+(n-m)|}$.

The first summation is zero, by eqn. (52).

The second and third summations may be shown to cancel out term by term.

(6.2.5) Case (e).

To prove

$$\sum_{k=2,4,\ldots}^{\infty} (r_{|k-m|} r_{|k+n|} + r_{|k+m|} r_{|k-n|}) + r_m r_n = 0 \quad (65)$$

for m and n odd, and $m + n \neq 0$.

The expression may be rewritten as

$$\sum_{M=1,3,\ldots}^{\infty} [r_M r_{|M+(m+n)|} + r_M r_{|M-(m+n)|}] + \sum_{M=2-m}^{-1} r_{|M|} r_{|M+(m+n)|} - \sum_{M=1}^{m-2} r_M r_{|M-(m+n)|} \quad (66)$$

The first summation is zero, by eqn. (52).

The second and third summations may be shown to cancel out term by term.

(6.2.6) Case (f).

To prove

$$\sum_{k=2,4,\ldots}^{\infty} (r_{|k-n|}^2 + r_{|k+n|}^2) + r_n^2 = \tfrac{1}{4}\pi^2 r_1^2 \quad . \quad . \quad (67)$$

The left-hand side may be rewritten as

$$2 \sum_{M=1,3,\ldots}^{\infty} r_M^2 = 2 \sum_{M=1,3,\ldots}^{\infty} \frac{r_1^2}{M^2} = \tfrac{1}{4}\pi^2 r_1^2$$

(6.3) Calculations for Section 1.3

Since it is always assumed that r(t) is a square-wave, the coefficients of the Fourier series, eqns. (2) and (3), can be expressed in terms of r_b and r_f:

$$r_0 = \tfrac{1}{2}(r_b + r_f)$$

$$r_1 = \frac{2}{\pi}(r_b - r_f)$$

Then, for the series modulator, Case 1, putting $R_{1+} = R_0 = R_S + R_R = R$ (say) in eqn. (18),

$$i_{1+} = \frac{-\tfrac{1}{2}V\frac{2}{\pi}(r_b - r_f)}{(R + \tfrac{1}{2}r_b + \tfrac{1}{2}r_f)^2 - \tfrac{1}{4}(r_b - r_f)^2} \quad . \quad . \quad (68)$$

and the conversion loss ratio $= \tfrac{1}{2}V/\tfrac{1}{2}i_{1+}R$, if it is assumed $R_S = R_R = \tfrac{1}{2}R$, is

$$= -\frac{(R + \tfrac{1}{2}r_b + \tfrac{1}{2}r_f)^2 - \tfrac{1}{4}(r_b - r_f)^2}{(r_b - r_f)R/\pi}$$

$$= -\frac{R^2 + R(r_b + r_f) + r_b r_f}{(r_b - r_f)R/\pi} \quad . \quad . \quad . \quad . \quad (69)$$

The loss is thus a minimum when $R = \sqrt{(r_b r_f)} = \rho$, i.e. when $R_S = R_R = \tfrac{1}{2}\rho$.

The minimum conversion loss ratio (ignoring polarity) is then

$$\frac{2\rho^2 + \rho(\rho n + \rho/n)}{\frac{1}{\pi}\rho(\rho n - \rho/n)} = \frac{2 + n + 1/n}{n - 1/n}\pi \quad . \quad . \quad (70)$$

and if $n \gg 1/n$, this ratio equals $\pi(1 + 2/n)$, which gives a conversion loss of very nearly $9{\cdot}92 + 17{\cdot}4/n$ dB.

Case 2 is calculated similarly, except that in eqn. (18) we substitute

$$R_{1+} = R_S + R_R = 2R, \; R_0 = R_S = R.$$

In Case 3 the substitution in eqn. (18) is $R_{1+} = R_0 = R$,

but since this time the output voltage is $i_{1+}R_R = i_{1+}R$ (instead of $\tfrac{1}{2}i_{1+}R$ as in Case 1), the conversion loss ratio is one-half of that of Case 1, and the minimum conversion loss is 6 dB less.

The working for the shunt modulators is similar, bearing in mind the equivalences of Figs. 2 and 3.

Case 4.—In the equivalent circuit of Fig. 3, this case corresponds to

$$Z = R_{S0} \text{ at frequency } \omega_q,$$
$$= R_{R1+} \text{ at frequency } \omega_p + \omega_q,$$
$$= \infty \text{ at all other frequencies.}$$

Eqns. (4) and (5) therefore become simply

$$V = i_0(R_{S0} + r_0) + \tfrac{1}{2}i_{1+}r_1 \quad . \quad . \quad . \quad . \quad (71)$$

$$0 = i_{1+}(R_{R1+} + r_0) + \tfrac{1}{2}i_0 r_1 \quad . \quad . \quad . \quad (72)$$

giving the solution for i_{1+} shown in eqn. (21).

Case 5.—In this case, R is $R_S + R_{R1+}$ at frequency $\omega_p + \omega_q$, but is R_S at all other frequencies. Since we assume there are no even-order currents, eqns. (4) and (5) give

$$V = i_0(R_S + r_0) + \tfrac{1}{2}\sum_{m=1,3,\ldots}(i_{m+} + i_{m-})r_m \quad . \quad (73)$$

$$0 = i_{1+}(R_S + R_{R1+} + r_0) + \tfrac{1}{2}i_0 r_1 \quad . \quad . \quad . \quad (74)$$

$$0 = i_{1-}(R_S + r_0) + \tfrac{1}{2}i_0 r_1 \quad . \quad . \quad . \quad . \quad . \quad (75)$$

$$0 = i_{3+}(R_S + r_0) + \tfrac{1}{2}i_0 r_3 \quad . \quad . \quad . \quad . \quad . \quad (76)$$

etc.

Substituting the latter equations into the first gives the solution for i_{1+} shown in eqn. (22).

(6.4) Practical Circuits which Approximate to the Cases in Table 3

Many practical circuits use filters or tuned circuits in the input and output of the modulator, and when the frequency separations are sufficient, or the cut-off of the filters is suffi-

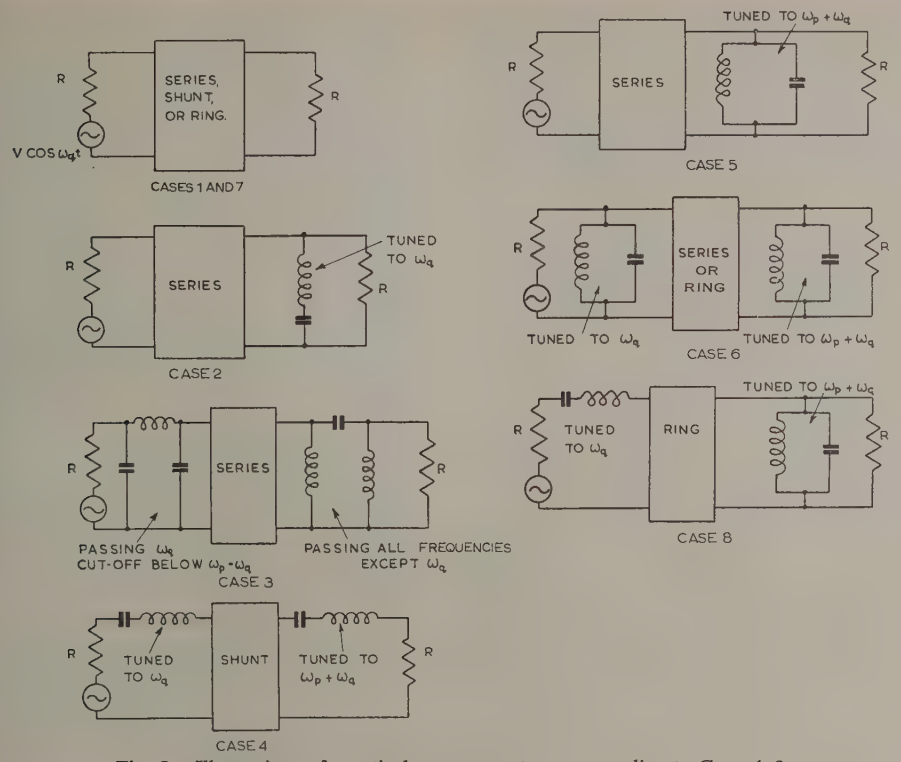


Fig. 8.—Illustrations of practical arrangements corresponding to Cases 1–8.

ciently sharp, or the Q-factors of the tuned circuits are sufficiently high, these arrangements generally correspond very closely to the cases tabulated in Table 3. A selection of such arrangements is shown in Fig. 8, and it will be seen that the results of the paper can be applied very widely in practice.

(6.5) Odd- and Even-Order Modulation Products in the Ring Modulator

A simple way to demonstrate that only odd-order currents flow in the output loop of the ring modulator, and only even-order currents in the input loop, is by the use of the compensation theorem. This theorem states that the voltages and currents in a circuit are not altered if a particular impedance Z, carrying a current I, is replaced by an e.m.f. generator of zero impedance and e.m.f. equal to $-IZ$. Thus in the single-loop modulator circuit of Fig. 3, the time-varying resistance r(t) can be replaced by an e.m.f. $-i.\mathrm{r}(t)$. Now the equations of this circuit are eqns. (4)–(6), and from these we deduce that the replacement e.m.f. can be represented as in Fig. 9, according to the following argument:

(a) It must be assumed that i_0 flows in the same direction as that in which V acts, as otherwise we could have a negative resistance at frequency ω_q.

(b) If, for the moment, it is assumed that there are no even-order currents other than that at frequency ω_q, then at frequency $n\omega_p \pm \omega_q$,

$$i_{n\pm}(R_{n\pm} + r_0) = -\tfrac{1}{2}i_0 r_n \text{ from eqn. (5)}$$

Therefore $i_{n\pm}$ flows in the opposite direction to i_0.

(N.B.—Throughout this reasoning, the directions refer to the direction of the current or e.m.f. as given by a particular expression. If, in this expression, any current or coefficient of r(t) is negative, then obviously the actual direction of the current or e.m.f. is reversed.)

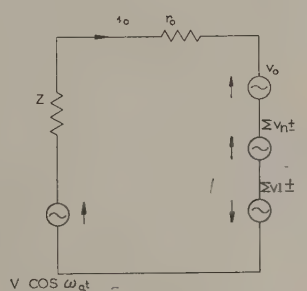


Fig. 9.—Equivalent single-loop circuit.

$$v_0 = \tfrac{1}{2}\sum_{m=1,3,\ldots}(i_{m+} + i_{m-})r_m \cos \omega_q t$$

$$\Sigma v_{n\pm} = \sum_{n=1\ 3,\ldots}\tfrac{1}{2}[i_0 r_n + \sum_{k=2,4,\ldots}(i_{k\pm}r_{|k-n|} + i_{k\mp}r_{|k+n|})]\cos(n\omega_p \pm \omega_q)t$$

$$\Sigma v_{l\pm} = \sum_{l=2,4,\ldots}\tfrac{1}{2}\sum_{m=1,3,\ldots}(i_{m\pm}r_{|m-l|} + i_{m\mp}r_{|m+l|})\cos(l\omega_p \pm \omega_q)t$$

(c) At frequency ω_q, therefore, eqn. (4) shows that there is a back e.m.f.

$$\tfrac{1}{2}\sum_{m=1,3,\ldots}(i_{m+} + i_{m-})r_m$$

(d) Now, if there are even-order currents, then at frequency $\omega_p \pm \omega_q$, from eqn. (6),

$$i_{l\pm}(R_{l\pm} + r_0) = -\tfrac{1}{2}\sum_{m=1,3,\ldots}(i_{m\pm}r_{|m-l|} + i_{m\pm}r_{|m+l|})$$

so that $i_{l\pm}$ flows in the opposite direction to the odd-order currents $i_{m\pm}$. This means that $i_{l\pm}$ flows in the same direction as i_0, so that the reasoning of (*b*) is still valid if even-order currents exist.

This representation of $-i.\mathrm{r}(t)$ can now be applied to the ring modulator circuit shown in Fig. 4B, noting that in $\mathrm{r}(t)$ all the coefficients r_n (except r_0) are reversed in sign relative to those

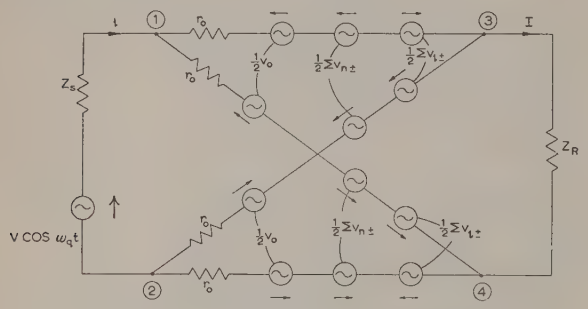


Fig. 10.—Equivalent circuit of ring modulator.

in $\mathrm{r}_+(t)$. The resultant circuit obtained is shown in Fig. 10, and the arguments are as follows:

(i) We note that the symmetry of the network necessitates that the magnitude of current at frequency ω_q in the cross-arms is the same as in the series arms. The direction of the currents must both be away from node 1, since if the directions were opposite at this point, Kirchhoff's current law would necessitate a value of zero for the current i_0 drawn from the source. This would mean that no energy could be used in the circuit and all currents would be zero. So, clearly, the current at frequency ω_q drawn from the source divides equally at node 1. Evidently, if the symbol i is used for the current in the input loop, the currents in the lattice arms are one-half of i, and all the e.m.f.'s require a coefficient $\frac{1}{2}$ if expressed in terms of i according to Fig. 9.

(ii) It is therefore clear that the back-e.m.f.'s $\frac{1}{2}v_0$ point towards node 1 and away from node 2, as shown in Fig. 10.

(iii) Since the back-e.m.f. at frequency ω_q acts in the same direction in both series and cross-arms, although r_m is reversed in sign in the cross-arm [by eqn. (39*b*)], it follows that the odd-order currents $i_{m\pm}$ are reversed in the cross arm relative to the series arm. This means the odd-order e.m.f.'s $(\frac{1}{2}v_{n\pm})$ are reversed as shown in Fig. 10.

(iv) In the expression for the e.m.f. $v_{l\pm}$ in the cross-arm, both $i_{m\pm}$ and $r_{|m\pm l|}$ are reversed in sign, and therefore the even-order e.m.f.'s act in the same direction in series and cross-arms, as shown in Fig. 10.

It is now evident that, at frequency ω_q and all even-order modulation-product frequencies, there is no p.d. between terminals 3 and 4, and no currents at these even-order frequencies flow in the output circuit. It is similarly evident that at all odd-order modulation-product frequencies there is no p.d. between terminals 1 and 2, and no odd-order currents flow in the input circuit.

It should be noted that this discussion has not assumed that $\mathrm{r}(t)$ is a square wave, but merely that it contains only odd-order harmonics.

[The discussion on the above paper will be found on page 281.]

521.376.234

The Institution of Electrical Engineers
Paper No. 3187 E
Jan. 1960

THE INPUT IMPEDANCE OF RECTIFIER MODULATORS

By Professor D. G. TUCKER, D.Sc., Member.

The paper was first received 23rd April, and in revised form 24th June, 1959. It was published in January, 1960, and was read before the ELECTRONICS AND COMMUNICATIONS SECTION *11th January, 1960.)*

SUMMARY

It is shown how the input impedance of a rectifier modulator of series, shunt or ring type can be calculated, and a large number of cases are tabulated, assuming (*a*) that the terminating impedances, while being as frequency selective as desired, are nevertheless purely resistive at all modulation-product frequencies where they are not zero or infinite, and (*b*) that the rectifiers have a square-wave variation of resistance with time. It is clearly shown that in all cases except one (i.e. a ring modulator with constant-resistance load) the input impedance—defined as the ratio of the input voltage to the input current at the applied signal frequency—has the peculiar property of being dependent on the nature and magnitude of the signal-source impedance from which it is being measured.

LIST OF SYMBOLS

Z = Impedance in single-loop modulator circuit representing terminating impedances.
Z_S = Signal-source impedance.
Z_R = Receiving (load) impedance.
Z_i = Input impedance of modulator.
Z_{i0}, Y_{i0} = Input impedance and admittance of modulator measured in terms of voltage and current at frequency of applied signal.
R, G = Values of Z, Y when non-reactive.
ω_q = Angular frequency of signal.
ω_p = Angular frequency of carrier (or local oscillator).
V' = E.M.F. of signal source in actual circuit.
V = E.M.F. of signal source in single-loop equivalent circuit.
I'_S = Current of Norton equivalent of signal source in actual circuit.
I_S = Current of Norton equivalent of signal source in single-loop equivalent circuit.
$\mathrm{r}(t)$ = Time-varying rectifier resistance.
$\mathrm{g}(t)$ = Time-varying rectifier conductance.
$\mathrm{v}(t)$ = Voltage across rectifier.
$\mathrm{i}(t)$ = Current through rectifier (and current in input loop of ring modulator).
$\mathrm{v}'(t)$ = Voltage across input terminals of modulator.
$\mathrm{i}'(t)$ = Current flowing into input terminals of modulator.
$\mathrm{I}(t)$ = Current in output loop of ring modulator.
n = Odd integer.
l = Even integer.

Subscripts:

0 = Value at frequency ω_q.
$1+$ = Value at frequency $\omega_p + \omega_q$.
$1-$ = Value at frequency $\omega_p - \omega_q$.
$n \pm$ = Value at frequency $n\omega_p \pm \omega_q$.

(1) INTRODUCTION

Calculations of the input impedance of rectifier modulators, terminated with circuits of various types of frequency dependence, have only once, to the author's knowledge, been previously published. The paper[1] drew attention to a very interesting and peculiar property of this input impedance, namely that it is dependent on the impedance of the signal circuit from which it is measured. No ordinary passive linear time-invariant circuit has this property. At the time of the previous paper, however, there were several aspects of the circuit calculations of rectifier modulators which were not properly understood, with the consequence that, while the general conclusions were correct, two of the individual cases which were considered were wrongly calculated. The major difficulties have now been resolved, however, and a recent paper[2] sets out a theoretical basis of modulator-circuit analysis, together with some conversion-loss calculations and experimental verification. It is the object of the present work to use this theory to obtain fresh calculations of the input impedance, and many useful cases are tabulated.

It is quite possible to calculate the input impedance of modulators when the circuit terminations have any prescribed variation with frequency, but the work is then very laborious. A very great simplification is obtained if it is assumed that the sending and receiving circuits have impedances which, while varying with frequency as much as desired, nevertheless have a purely resistive impedance at all the modulation-product frequencies where they are not zero or infinite. This simplification will be assumed throughout the paper; it is not at all an unpractical simplification, as the majority of practical cases approximate (or can easily be made to approximate) to this condition.

Other assumptions made for simplicity are:

(*a*) The rectifiers are either non-reactive or their reactance is absorbed into the terminations (i.e. into the external circuits), and they are switched by the carrier (or local oscillator) at zero voltage between a constant back resistance, r_b, and a constant forward resistance, rf, so that the time variation of the rectifier resistance, $\mathrm{r}(t)$, is a square-wave: thus

$$\mathrm{r}(t) = r_0 + \sum_{n=1,3,\ldots}^{\infty} r_n \cos n\omega_p t \quad . \quad . \quad . \quad . \quad (1)$$

where $$r_n = (-1)^{(n-1)/2} r_1/n \quad . \quad . \quad . \quad . \quad . \quad (2)$$

and n is odd.

Clearly $$r_0 = \tfrac{1}{2}(r_b + r_f) \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

and $$r_1 = \frac{2}{\pi}(r_b - r_f) \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Alternatively, conductances may be used; thus

$$\mathrm{g}(t) = \frac{1}{\mathrm{r}(t)} = g_0 + \sum_{n=1,3,\ldots}^{\infty} g_n \cos n\omega_p t \quad . \quad . \quad (5)$$

where $$g_n = (-1)^{(n-1)/2} g_1/n \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Also $$g_0 = \frac{1}{2}\left(\frac{1}{r_f} + \frac{1}{r_b}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

and $$g_1 = \frac{2}{\pi}\left(\frac{1}{r_f} - \frac{1}{r_b}\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

Professor Tucker is Professor of Electrical Engineering, University of Birmingham.

(*b*) The circuit is linear, i.e. the values of the circuit-elements, including $r(t)$, are independent of the signal voltage and current, and they are switched by the carrier (or local oscillator) at zero signal voltage or current. In practice, this condition is achieved by making the signal voltages and currents much smaller than those of the carrier.

(2) CIRCUIT ARRANGEMENTS, DEFINITIONS AND EQUATIONS

The modulator circuits which are considered are the well-known series, shunt and ring modulators,[3] shown in Figs. 1–3.

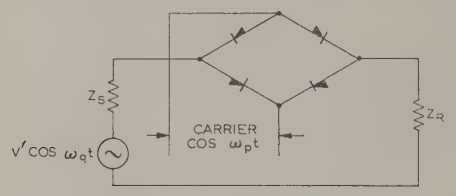


Fig. 1.—Series modulator.

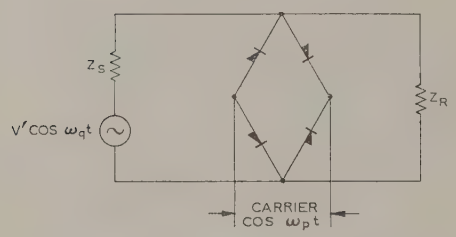


Fig. 2.—Shunt modulator.

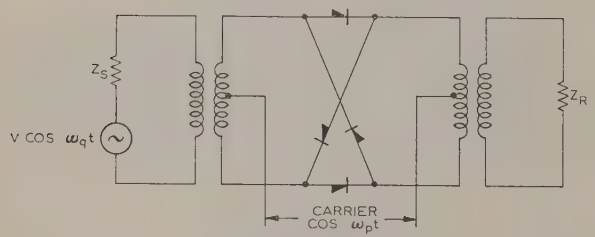


Fig. 3.—Ring modulator.

Their effective transmission circuits are shown in Figs. 4–6, respectively.

The input impedance discussed in the paper is that measured at the pair of terminals to which the input signal is applied, as shown in Figs. 4–6. This impedance is represented generally by Z_i. There is no doubt that if the full time-varying currents and voltages are considered, then if, for instance, we define

$$Z_i(t) = v'(t)/i'(t) \quad . \quad . \quad . \quad . \quad . \quad (9)$$

this impedance is independent of the signal-source impedance. But it varies with time, generally in a very complex manner, and for design purposes, e.g. in planning a telephone system, it is useless. What would be measured in practice, according to any ordinary method, is the component of $Z_i(t)$ at the frequency of the applied signal (i.e. at angular frequency ω_q). This is a

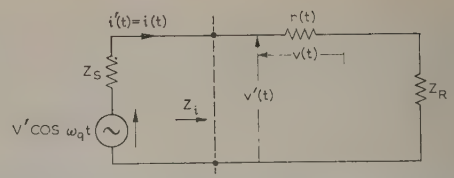


Fig. 4.—Effective transmission circuit of series modulator.

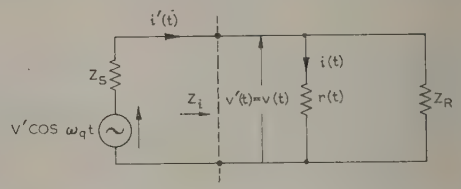


Fig. 5.—Effective transmission circuit of shunt modulator.

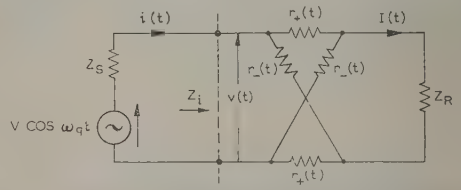


Fig. 6.—Effective transmission circuit of ring modulator.

$$r_+(t) = r_0 + \sum_{n=1,3,\ldots}^{\infty} r_n \cos n\omega_p t$$

$$r_-(t) = r_0 - \sum_{n=1,3,\ldots}^{\infty} r_n \cos n\omega_p t$$

steady impedance, Z_{i0}, which is clearly the ratio of the components of $v'(t)$ and $i'(t)$ at frequency ω_q; i.e.

$$Z_{i0} = v_0'/i_0' \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

It is this impedance which has the peculiar property of being dependent on Z_S and which is the subject of the present study.

It was shown in Reference 2 that with the simplifications discussed in Section 1 the three types of modulator can all be represented by the same set of circuit equations. These equations comprise one for each frequency ($n\omega_p \pm \omega_q$ or $l\omega_p \pm \omega_q$, where n is odd and l is even) which can exist in the circuit. Since, in the ring modulator,[2] only currents and voltages of even-order modulation products exist in the input loop and only those of odd-order products in the output loop, it is convenient in setting out the equations to use i for even-order currents and I for odd-order currents in the series and shunt modulators. It is convenient, too, to reduce both series and shunt modulators to

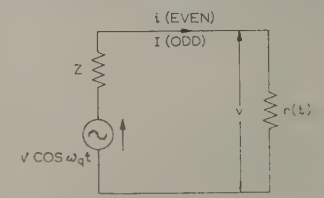


Fig. 7.—Equivalent single-loop modulator circuit.

the same single-loop circuit (as shown in Fig. 7) by combining the sending and receiving terminations, so that, in the series modulator,

$$V = V' \quad Z = Z_S + Z_R \quad . \quad . \quad . \quad . \quad . \quad (11)$$

Table 1

SERIES MODULATOR

Case No.	Sending-end impedance, Z_S	Receiving-end impedance, Z_R	Input impedance at signal frequency, Z_{i0}	Numerical example
1	Constant resistance R_S	Constant resistance R_R	$R_R + r_0 - \frac{\pi^2}{16}\frac{r_1^2}{R_S + R_R + r_0}$	ohms 3 700
2	Constant resistance R_S	Constant resistance R_R, except zero at frequency ω_q	$r_0 - \frac{\pi^2}{16}\frac{r_1^2}{R_S + R_R + r_0}$	2 500
3	Constant resistance R_S	R_{R1+} at $\omega_p + \omega_q$ Zero at all other frequencies	$\frac{r_0 - R_S D^*}{1 + D}$	1 950
4	R_{S0} at ω_q Zero at $\omega_p + \omega_q$ ∞ at all other frequencies	Constant resistance R_R	$R_R + r_0 - \frac{\frac{1}{4}r_1^2}{R_R + r_0}$	16 800
5	Constant resistance R_S	R_{R1+} at $\omega_p + \omega_q$ Zero at ω_q ∞ at all other frequencies	$r_0 - \frac{\frac{1}{4}r_1^2}{R_S + R_{R1+} + r_0}$	16 000
6	R_{S0} at ω_q Zero at all other frequencies	Constant resistance R_R	$R_R + r_0 - \frac{\pi^2}{16}\frac{r_1^2}{R_R + r_0}$	2 800
7	R_{S0} at ω_q Zero at $\omega_p + \omega_q$ and at all other frequencies either (*a*) ∞ or (*b*) any value	R_{R1+} at $\omega_p + \omega_q$ Zero at ω_q and at all other frequencies either (*a*) any value or (*b*) ∞	$r_0 - \frac{\frac{1}{4}r_1^2}{R_{R1+} + r_0}$	15 600
8	R_{S0} at ω_q Zero at all other frequencies	R_{R1+} at $\omega_p + \omega_q$ Zero at all other frequencies	$\frac{1 + g_0 R_{R1+}}{g_0 + R_{R1+}(g_0^2 - \frac{1}{4}g_1^2)}$	700

$$* D = \frac{\frac{1}{4}r_1^2}{(R_S + r_0)(R_S + R_{R1+} + r_0) - \frac{\pi^2}{16}r_1^2} + \frac{\frac{1}{4}r_1^2\left(\frac{\pi^2}{4} - 1\right)}{(R_S + r_0)^2 - \frac{\pi^2}{16}r_1^2}$$

and in the shunt modulator

$$V = V'\frac{Z_R}{Z_S + Z_R} \qquad Z = \frac{Z_S Z_R}{Z_S + Z_R} \quad . \quad . \quad (12)$$

The set of equations is then,

at frequency ω_q,

$$V = i_0(R_0 + r_0) + \tfrac{1}{2}\sum_{m=1,3,\ldots}(I_{m+} + I_{m-})r_m \quad . \quad (13)$$

at frequency $n\omega_p \pm \omega_q$,

$$0 = I_{n\pm}(R_{n\pm} + r_0) + \tfrac{1}{2}i_0 r_n + \tfrac{1}{2}\sum_{k=2,4,\ldots}(i_{k\pm}r_{|k-n|} + i_{k\mp}r_{k+n}) \quad . \quad . \quad . \quad . \quad (14)$$

and at frequency $l\omega_p \pm \omega_q$,

$$0 = i_{l\pm}(R_{l\pm} + r_0) + \tfrac{1}{2}\sum_{m=1,3,\ldots}(I_{m\pm}r_{|m-l|} + I_{m\mp}r_{m+l}) \quad (15)$$

The symbol R is used now instead of Z since the values are purely resistive at all modulation-product frequencies. Evidently R is a combination of Z_S and Z_R in series and shunt modulators; but in the ring modulator, when R has an even subscript it refers to Z_S, and when it has an odd subscript it refers to Z_R.

A completely dual set of equations may be obtained if voltages and admittances are used on the right-hand side of the equations, with current sources on the left-hand side. This set is advantageous in leading to simple solutions in certain circumstances (e.g. when Z is zero at some frequencies) where eqns. (13)–(15) are difficult to solve; and, conversely, there are circumstances (e.g. when Z is infinite at some frequencies) where the first set is preferable. The dual set is evidently,

at frequency ω_q,

$$I_S = v_0(G_0 + g_0) + \tfrac{1}{2}\sum_{m=1,3,\ldots}(V_{m+} + V_{m-})g_m \quad . \quad (16)$$

at frequency $n\omega_p \pm \omega_q$,

$$0 = V_{n\pm}(G_{n\pm} + g_0) + \tfrac{1}{2}v_0 g_n + \tfrac{1}{2}\sum_{k=2,4,\ldots}(v_{k\pm}g_{|k-n|} + v_{k\mp}g_{k+n}) \quad . \quad . \quad . \quad . \quad (17)$$

and at frequency $l\omega_p \pm \omega_q$,

$$0 = v_{l\pm}(G_{l\pm} + g_0) + \tfrac{1}{2}\sum_{m=1,3,\ldots}(V_{m\pm}g_{|m-l|} + V_{m\mp}g_{m+l}) \quad (18)$$

where v, or V with a subscript, is the voltage across r(t) in Fig. 7.

(3) CALCULATION OF INPUT IMPEDANCE

The input impedance Z_{i0} may usually be obtained from these equations in the form

$$Z_{i0} = \frac{V' - i_0' R_{S0}}{i_0'} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

or

$$Y_{i0} = \frac{1}{Z_{i0}} = \frac{I_S' - v_0' G_{S0}}{v_0'} \quad . \quad . \quad . \quad . \quad (20)$$

Table 2

SHUNT MODULATOR

Case No.	Sending-end impedance, Z_S	Receiving-end impedance, Z_R	Input impedance at signal frequency, Z_{i0}	Numerical example
				ohms
1	Constant resistance R_S	Constant resistance R_R	$\dfrac{R_R\left[r_0 - \dfrac{\pi^2}{16}\dfrac{r_1^2}{\dfrac{R_S R_R}{R_S+R_R}+r_0}\right]}{R_R + r_0 - \dfrac{\pi^2}{16}\dfrac{r_1^2}{\dfrac{R_S R_R}{R_S+R_R}+r_0}}$	570
2	Constant resistance R_S	Constant resistance R_R, except ∞ at frequency ω_q	$r_0 - \dfrac{\pi^2}{16}\dfrac{r_1^2}{\dfrac{R_S R_R}{R_S+R_R}+r_0}$	1 100
3	Constant resistance R_S	R_{R1+} at $\omega_p+\omega_q$ ∞ at all other frequencies	$\dfrac{r_0 - R_S A^*}{1+A}$	1 380
4	R_{S0} at ω_q ∞ at all other frequencies	Constant resistance R_R	$\dfrac{R_R\left[r_0 - \dfrac{\pi^2}{16}\dfrac{r_1^2}{R_R+r_0}\right]}{R_R + r_0 - \dfrac{\pi^2}{16}\dfrac{r_1^2}{R_R+r_0}}$	700
5	Constant resistance R_S	R_{R1+} at $\omega_p+\omega_q$ ∞ at ω_q Zero at all other frequencies	$\dfrac{1+g_0 R_S R_{R1+}/(R_S+R_{R1+})}{g_0+(g_0^2-\frac{1}{4}g_1^2)R_S R_{R1+}/(R_S+R_{R1+})}$	640
6	R_{S0} at ω_q ∞ at $\omega_p+\omega_q$ Zero at all other frequencies	Constant resistance R_R	$\dfrac{1/R_R+g_0}{(1/R_R+g_0)^2-\frac{1}{4}g_1^2}$	440
7	R_{S0} at ω_q ∞ at all other frequencies	R_{R1+} at $\omega_p+\omega_q$ ∞ at all other frequencies	$r_0 - \dfrac{\frac{1}{4}r_1^2}{(R_{R1+}+r_0}$	15 600
8	R_{S0} at ω_q ∞ at $\omega_p+\omega_q$ and at all other frequencies either (*a*) zero or (*b*) any value	R_{R1+} at frequency $\omega_p+\omega_q$ ∞ at ω_q and at all other frequencies either (*a*) any value or (*b*) zero	$\dfrac{1+g_0 R_{R1+}}{g_0+R_{R1+}(g_0^2-\frac{1}{4}g_1^2)}$	700

$$* A = \frac{\frac{1}{4}r_1^2}{(R_S+r_0)\left(\frac{R_S R_{R1+}}{R_S+R_{R1+}}+r_0\right)-\frac{\pi^2}{16}r_1^2} + \frac{\frac{1}{4}r_1^2\left(\frac{\pi^2}{4}-1\right)}{(R_S+r_0)^2-\frac{\pi^2}{16}r_1^2}$$

where R_{S0} and G_{S0} are the values of R_S and $G_S(=1/R_S)$ at the applied frequency ω_q.

A simple case will now be worked out to illustrate the process and the effects.

Consider a shunt modulator (Fig. 5) in which Z_S is a constant resistance R_S at all frequencies and Z_R is a constant resistance R_R at all frequencies except that it is infinite at the signal frequency ω_q. Then at all frequencies except ω_q the combined source and load impedance (Z in Fig. 7) is a constant resistance $R_S R_R/(R_S+R_R)$. It is shown in Reference 2 that if the impedance Z is a constant resistance at all odd-order frequencies, then no even-order currents flow. Consequently all even-order currents are zero in the present case, and eqn. (13) becomes

$$V = i_0(R_S+r_0) + \tfrac{1}{2}i_{1+}r_1 + \tfrac{1}{2}i_{1-}r_1 + \tfrac{1}{2}i_{3+}r_3 + \ldots \quad (21)$$

while eqn. (14) becomes

$$0 = i_{n\pm}\left(\frac{R_S R_R}{R_S+R_R}+r_0\right) + \tfrac{1}{2}i_0 r_n \quad . \quad . \quad . \quad (22)$$

Eqn. (22) can be directly substituted into eqn. (21) for all values of n (odd) giving, since $V = V'$ in this case,

$$Z_{i0} = \frac{V' - i_0' R_S}{i_0'} = r_0 - \frac{1}{2}\,\frac{r_1^2 + r_3^2 + \ldots}{\dfrac{R_S R_R}{R_S+R_R}+r_0} \quad . \quad (23)$$

$$= r_0 - \frac{\pi^2}{16}\,\frac{r_1^2}{\dfrac{R_S R_R}{R_S+R_R}+r_0} \quad (24)$$

since if r(t) is a square-wave, then

$$r_1^2 + r_3^2 + r_5^2 + \ldots = r_1^2\left(1+\frac{1}{3^2}+\frac{1}{5^2}+\ldots\right) = \frac{\pi^2}{8}r_1^2 \quad (25)$$

Note that in this case, since $R_R = \infty$ at frequency ω_q, then $i_0' = i_0$.

It is clear from this that Z_{i0} is dependent on R_S. In other words, the ratio v_0'/i_0' at frequency ω_q is not a single-valued quantity but depends on the impedance of the source from which it is measured. This might almost be thought to be at variance

Table 3

RING MODULATOR

Case No.	Sending-end impedance, Z_S	Receiving-end impedance, Z_R	Input impedance at signal frequency, Z_{i0}	Numerical example
1	Any value	Constant resistance R_R	$r_0 - \frac{\pi^2}{16}\frac{r_1^2}{R_R + r_0}$	ohms 1 640
2	Constant resistance R_S	R_{R1+} at $\omega_p + \omega_q$ ∞ at all other frequencies	$\frac{r_0 - R_S r_1^2/4B^*}{1 + r_1^2/4B}$	4 800
3	Constant resistance R_S	R_{R1+} at $\omega_p + \omega_q$ Zero at all other frequencies	$1 \Big/ \frac{g_0 - \frac{1}{R_S}\frac{\frac{1}{4}g_1^2}{C}\dagger}{1 + \frac{1}{4}g_1^2/C}$	750
4	R_{S0} at ω_q ∞ at all other frequencies	R_{R1+} at $\omega_p + \omega_q$ ∞ at all other frequencies	$r_0 - \frac{\frac{1}{4}r_1^2}{(R_{R1+} + r_0)}$	15 600
5	R_{S0} at ω_q Zero at all other frequencies	R_{R1+} at $\omega_p + \omega_q$ Zero at all other frequencies	$\frac{1 + g_0 R_{R1+}}{g_0 + R_{R1+}(g_0^2 - \frac{1}{4}g_1^2)}$	700
6	R_{S0} at ω_q ∞ at all other frequencies	R_{R1+} at $\omega_p + \omega_q$ Zero at all other frequencies	$r_0 - \frac{\frac{1}{4}r_1^2}{R_{R1+} + r_0} - \frac{1}{4}\left(\frac{\pi^2}{4} - 1\right)\frac{r_1^2}{r_0}$	920
7	R_{S0} at ω_q Zero at all other frequencies	R_{R1+} at $\omega_p + \omega_q$ ∞ at all other frequencies	$1 \Big/ \left[g_0 - \frac{\frac{1}{4}g_1^2}{\frac{1}{R_{R1+}} + g_0} - \frac{1}{4}\left(\frac{\pi^2}{4} - 1\right)\frac{g_1^2}{g_0}\right]$	3 700

* $B = (R_{R1+} + r_0)(R_S + r_0) - \frac{\pi^2}{16}r_1^2$

† $C = \left(\frac{1}{R_{R1+}} + g_0\right)\left(\frac{1}{R_S} + g_0\right) - \frac{\pi^2}{16}g_1^2$

with Ohm's law. Nevertheless the physical explanation of the effect is simple enough. R_S forms part of the load termination (effectively) for modulation-product frequencies, so that naturally the input impedance depends on its value.

It still remains to determine whether Z_{i0} depends on the value of R_S at the applied frequency ω_q. Thus the value of R_S at this frequency is distinguished by writing it as R_{S0}. Then eqn. (21) becomes

$$V = i_0(R_{S0} + r_0) + \tfrac{1}{2}i_{1+}r_1 + \tfrac{1}{2}i_{1-}r_1 + \tfrac{1}{2}i_{3+}r_3 + \ldots \quad (26)$$

but eqn. (22) remains as before, with R_S. Thus

$$Z_{i0} = \frac{V' - i_0' R_{S0}}{i_0'} = r_0 - \frac{1}{2}\frac{r_1^2 + r_3^2 + \ldots}{\frac{R_S R_R}{R_S + R_R} + r_0} \quad . \quad (27)$$

as before, and evidently Z_{i0} is quite independent of R_{S0}.

A large number of arrangements of the terminating impedances has been analysed, and the input impedances are tabulated in Tables 1–3 for series, shunt and ring modulators, respectively. An outline of the calculations for shunt and ring modulators is given in the Appendices. The calculations for the series modulator are not given, as it is a little-used circuit and the working is similar to that of the shunt modulator. The results show the dependence of Z_{i0} on Z_S in all cases* except one, namely the ring modulator with a constant-resistance receiving termination, which is discussed below. Moreover, in all cases Z_{i0} is independent of R_{S0}, the value of Z_S at the applied frequency ω_q.

To illustrate the magnitudes involved in this matter, a numerical example has been included for each case, based on a modulator working with sending and receiving terminations each of 1 200 ohms at the frequency of applied signal and wanted sideband, respectively. The rectifier resistances are taken as $r_b = 50\,000$ ohms and $r_f = 250$ ohms.

It can be seen how greatly the input impedance varies from one condition to another. As a further illustration of the effects, the dependence of Z_{i0} on R_S has been shown graphically for a few suitable cases in Fig. 8, where the same numerical values have been taken except for regarding R_S as a variable.

The ring modulator with Z_R a constant resistance (strictly required to be a constant resistance only at all odd-order frequencies, since no others are present in the output loop of a ring modulator) is a special case because, if r(t) is a square-wave, the rectifiers merely commutate the connection to the load, which cannot store energy, being a pure resistance. Consequently the input impedance is a constant resistance not varying with time or frequency. Clearly, therefore, the applied e.m.f. at frequency ω_q cannot generate currents in the input loop at any other frequencies, and only the frequency ω_q exists in this loop. This means that Z_{i0} cannot be affected by the value of Z_S at any frequency.

The paper is, strictly, concerned only with modulators where r(t) is a square wave (although all the processes of calculation are valid if r(t) is not a square wave, provided that it is sym-

* Z_S (or R_S) does not appear explicitly in all the formulae, however; this is because its value at frequencies other than ω_q has been specified as zero or infinity. An example is shunt case No. 4, which may be derived from case No. 1 by putting $Z_S = \infty$ except at ω_q and so replacing R_S in the formula by ∞. This can be done, however, only with the knowledge that R_{S0} does not affect Z_{i0}, and it would be unhelpful to a casual user of the Tables not to show cases like shunt case No. 4 separately.

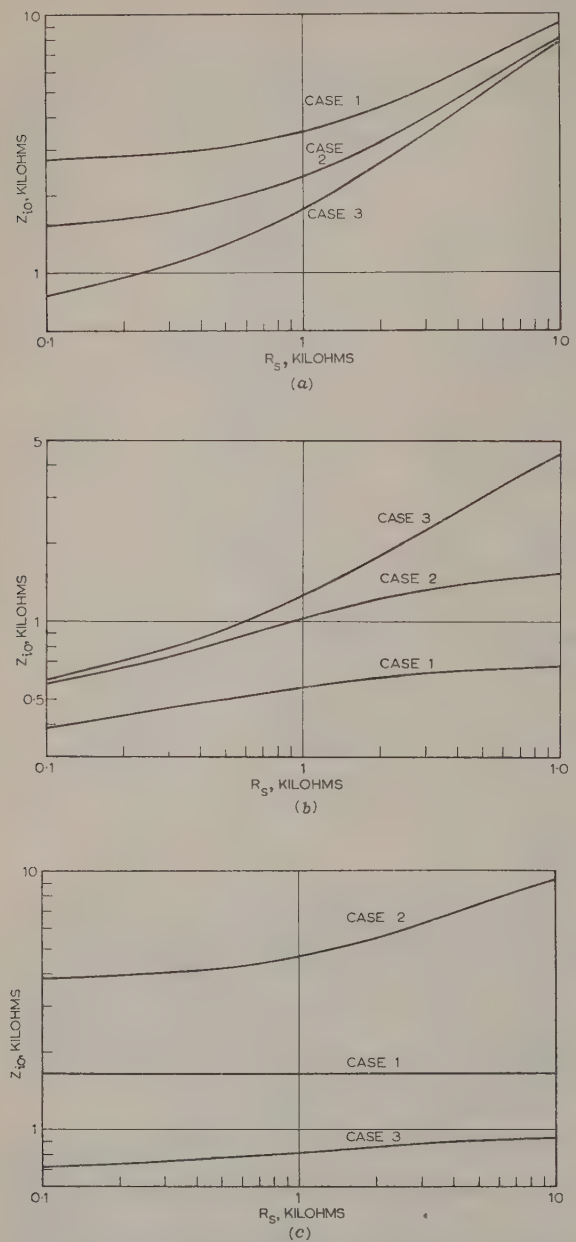


Fig. 8.—Modulator input impedance, Z_{i0}, versus signal-source resistance, R_S, for various modulator arrangements.

(*a*) Series modulator.
(*b*) Shunt modulator.
(*c*) Ring modulator.

metrical and thus contains only odd harmonics of ω_p). Nevertheless, it is of some interest and practical importance to point out that a constant-resistance input can be obtained in a ring modulator when $r(t)$ is not only not square but is not symmetrical. It was shown by the author many years ago[4] that if rectifiers with an exponential resistance/voltage relationship were used in a ring modulator, iterative network theory could be applied to obtain a constant input resistance. Thus, it is well known that if a symmetrical lattice network has series arms Z_1 and crossed arms Z_2, it is iterative with respect to a load impedance Z_R if

$$Z_1 Z_2 = Z_R^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (28)$$

Therefore the input impedance is a constant resistance equal to the load resistance when the network is terminated in a constant pure resistance, R_R, and

$$Z_1 Z_2 = R_R^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)$$

Now, if in the ring modulator Z_1 and Z_2 are the rectifier resistances,

$$Z_1 = r_+(t) = a\varepsilon^{-bV(t)} \quad . \quad . \quad . \quad . \quad (30)$$

and

$$Z_2 = r_-(t) = a\varepsilon^{+bV(t)} \quad . \quad . \quad . \quad . \quad (31)$$

where $V(t)$ is the longitudinally applied carrier voltage, then

$$Z_1 Z_2 = a^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (32)$$

so that a constant input resistance, independent of time and frequency, is obtained if

$$R_R = a \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (33)$$

If there is a constant resistance (e.g. the spreading resistance of the rectifiers) added to eqns. (30) and (31), this is easily allowed for by transferring it to the terminations by the well-known lattice-network theorem.

(4) CONCLUSIONS

It has been shown that the input impedance of a rectifier modulator is, in general, dependent on the signal source impedance from which it is measured. Results of practical interest have been tabulated.

(5) REFERENCES

(1) TUCKER, D. G.: 'Two Notes on the Performance of Rectifier Modulators, Part 1: The Input Impedance of Rectifier Modulators with Frequency-Selective Terminations', *Proceedings I.E.E.*, Paper No. 1384 R, November, 1952 (**99**, Part III, p. 400).

(2) HOWSON, D. P., and TUCKER, D. G.: 'Rectifier Modulators with Frequency-Selective Terminations, with particular reference to the Effect of Even-Order Modulation Products' (see page 261).

(3) TUCKER, D. G.: 'Modulators and Frequency-Changers' (Macdonald, 1953).

(4) TUCKER, D. G.: 'Some Aspects of the Design of Balanced Rectifier Modulators for Precision Applications', *Journal I.E.E.*, 1948, **95**, Part III, p. 161.

(6) APPENDICES

(6.1) Calculations for Table 2: Shunt Modulator

Case 1.—R_S and R_R both constant resistances. Let $R_S R_R/(R_S + R_R) = R'$. This is the effective impedance, Z, for all frequencies; as it is a constant resistance, then (by Reference 2) no even-order currents flow. Therefore eqn. (13) becomes

$$V = i_0(R' + r_0) + \tfrac{1}{2}i_{1+}r_1 + \tfrac{1}{2}i_{1-}r_1 + \tfrac{1}{2}i_{3+}r_3 \ldots \quad (34)$$

and eqn. (14) becomes

$$0 = i_{n\pm}(R' + r_0) + \tfrac{1}{2}i_0 r_n \quad . \quad . \quad . \quad (35)$$

Substituting eqn. (35) into eqn. (34),

$$V = i_0\left(R' + r_0 - \frac{1}{2}\frac{r_1^2 + r_3^2 + \ldots}{R' + r_0}\right) \quad . \quad . \quad . \quad (36)$$

low $Z_{i0} = R_R$ in parallel with v_0/i_0 (see Fig. 5), and, from ig. 7,

$$\frac{v_0}{i_0} = \frac{V - i_0 R'}{i_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (37)$$

o that

$$Z_{i0} = \frac{R_R\left(r_0 - \frac{\pi^2}{16}\frac{r_1^2}{R' + r_0}\right)}{R_R + r_0 - \frac{\pi^2}{16}\frac{r_1^2}{R' + r_0}} \quad . \quad . \quad . \quad (38)$$

sing eqn. (25).

Case 2.—Calculated in Section 3.

Case 3.—R_S a constant resistance, and $R_R = R_{R1+}$ at wanted ideband frequency, but infinite at all other frequencies. (The anted sideband frequency is always assumed to be $\omega_p + \omega_q$, ut the same results are obtained if it is $\omega_p - \omega_q$).

Since, in this case, the effective impedance is not a constant esistance for all odd-order frequencies, even-order currents xist. The calculation is then not as simple as in the previous ases. But Z is a constant resistance for all even-order fre- uencies. Thus, it can be seen from eqn. (30) of Reference 2 when R_0 is not assumed equal to R_l) that eqn. (14) of the present aper becomes

$$= i_{n\pm}(R_{n\pm} + r_0) + \tfrac{1}{2} i_0 r_n - \frac{1}{4(R_l + r_0)}\left[2 i_0 r_n (R_0 + r_0) - 2 V r_n + \frac{\pi^2}{4} i_{n\pm} r_1^2\right] \quad (39)$$

where it is assumed that Z is a constant resistance R_l for all even-order frequencies, i.e.

$$i_{n\pm}\left[(R_{n\pm} + r_0)(R_l + r_0) - \frac{\pi^2}{16} r_1^2\right] = \tfrac{1}{2} i_0 r_n (R_0 - R_l) - \tfrac{1}{2} V r_n \quad (40)$$

Therefore, substituting in eqn. (13),

$$V = i_0(R_{S0} + r_0) + \frac{\frac{1}{4} r_1^2 [i_0(R_{S0} - R_S) - V]}{(R' + r_0)(R_S + r_0) - \frac{\pi^2}{16} r_1^2} + \frac{\frac{1}{4} r_1^2 \left(\frac{\pi^2}{4} - 1\right)[i_0(R_{S0} - R_S) - V]}{(R_S + r_0)^2 - \frac{\pi^2}{16} r_1^2} \quad (41)$$

where $R' = \dfrac{R_{R1+} R_S}{R_{R1+} + R_S}$ and R_{S0} is distinguished from R_S merely to show how Z_{i0}, while being dependent on R_S, is not dependent on the value at the applied frequency ω_q. Now, since Z_R is infinite at frequency ω_q, $V = V'$, $v_0 = v_0'$, $i_0 = i_0'$, and therefore $Z_{i0} = (V - i_0 R_{S0})/i_0$ which, after reducing eqn. (41), becomes

$$Z_{i0} = \frac{r_0 - R_S A}{1 + A} \quad . \quad . \quad . \quad . \quad . \quad . \quad (42)$$

where

$$A = \frac{\frac{1}{4} r_1^2}{(R_S + r_0)(R' + r_0) - \frac{\pi^2}{16} r_1^2} + \frac{\frac{1}{4} r_1^2 \left(\frac{\pi^2}{4} - 1\right)}{(R_S + r_0)^2 - \frac{\pi^2}{16} r_1^2} \quad (43)$$

Case 4.—$R_S = R_{S0}$ at applied frequency ω_q, but infinity at all sideband frequencies, and R_R a constant resistance.

In this case there are no even-order currents, and eqn. (13) becomes, if $R_{S0} R_R/(R_{S0} + R_R) = R'$,

$$V = i_0(R' + r_0) + \tfrac{1}{2} i_{1+} r_1 + \tfrac{1}{2} i_{1-} r_1 + \tfrac{1}{2} i_{3+} r_3 + \ldots \quad (44)$$

and eqn. (14) becomes

$$0 = i_{n\pm}(R_R + r_0) + \tfrac{1}{2} i_0 r_n \quad . \quad . \quad . \quad . \quad (45)$$

Substituting eqn. (45) into eqn. (44),

$$V = i_0\left(R' + r_0 - \frac{\pi^2}{16}\frac{r_1^2}{R_R + r_0}\right) \quad . \quad . \quad . \quad (46)$$

Now $Z_{i0} = R_R$ in parallel with v_0/i_0, as in case 1, i.e.

$$Z_{i0} = \frac{R_R\left(r_0 - \frac{\pi^2}{16}\frac{r_1^2}{R_R + r_0}\right)}{R_R + r_0 - \frac{\pi^2}{16}\frac{r_1^2}{R_R + r_0}} \quad . \quad . \quad . \quad (47)$$

Case 5.—R_S is a constant resistance, and $R_R = R_{R1+}$ at wanted sideband frequency, but infinite at frequency ω_q and zero at all other frequencies.

In this case, the effective impedance is zero at all frequencies other than ω_q and $\omega_p + \omega_q$; at ω_q it is R_S and at $\omega_p + \omega_q$ it is $R' = R_S R_{R1+}/(R_S + R_{R1+})$. It pays, this time, to use eqns. (16)–(18). Thus, from eqn. (16),

$$\frac{V}{R_S} = v_0\left(\frac{1}{R_S} + g_0\right) + \tfrac{1}{2} v_{1+} g_1 \quad . \quad . \quad . \quad (48)$$

and from eqn. (17)

$$0 = v_{1+}\left(\frac{1}{R'} + g_0\right) + \tfrac{1}{2} v_0 g_1 \quad . \quad . \quad . \quad . \quad (49)$$

since all voltages are zero except v_0 and v_{1+}. Substituting eqn. (49) into eqn. (48),

$$\frac{V}{R_S} = v_0\left(\frac{1}{R_S} + g_0 - \frac{\frac{1}{4} g_1^2}{\frac{1}{R'} + g_0}\right) \quad . \quad . \quad . \quad (50)$$

Now applying eqn. (20), noting that $I_S' = V/R_S$, the input admittance is

$$Y_{i0} = g_0 - \frac{\frac{1}{4} g_1^2}{\frac{1}{R'} + g_0} \quad . \quad . \quad . \quad . \quad (51)$$

Therefore

$$Z_{i0} = \frac{1 + g_0 R'}{g_0 + R'(g_0^2 - \frac{1}{4} g_1^2)} \quad . \quad . \quad . \quad . \quad (52)$$

Case 6.—$R_S = R_{S0}$ at frequency ω_q, but infinite at wanted sideband frequency, and zero at all other frequencies; R_R a constant resistance.

This is similar to case 5, but the effective impedance is now $R' = R_{S0} R_R/(R_{S0} + R_R)$ at frequency ω_q and R_R at frequency $\omega_p + \omega_q$.

Therefore

$$\frac{V}{R'} = v_0\left(\frac{1}{R'} + g_0 - \frac{\frac{1}{4} g_1^2}{\frac{1}{R_R} + g_0}\right) \quad . \quad . \quad . \quad (53)$$

Now

$$Y_{i0} = \frac{1}{R_R} + \frac{i_0}{v_0} \quad . \quad . \quad . \quad . \quad . \quad (54)$$

and

$$i_0 = \frac{V}{R'} - \frac{v_0}{R'} \quad . \quad . \quad . \quad . \quad . \quad (55)$$

so that

$$\frac{i_0}{v_0} = g_0 - \frac{\frac{1}{4} g_1^2}{\frac{1}{R_R} + g_0} \quad . \quad . \quad . \quad . \quad . \quad (56)$$

Therefore $$Y_{i0} = \frac{1}{R_R} + g_0 - \frac{\frac{1}{4}g_1^2}{\frac{1}{R_R} + g_0} \quad . \quad . \quad . \quad . \quad (57)$$

and $$Z_{i0} = \frac{\frac{1}{R_R} + g_0}{\left(\frac{1}{R_R} + g_0\right)^2 - \frac{1}{4}g_1^2} \quad . \quad . \quad . \quad (58)$$

Case 7.—$R_S = R_{S0}$ at frequency ω_q, but infinite at all other frequencies, and $R_R = R_{R1+}$ at wanted sideband, but infinite at all other frequencies.

The equations here are simply

$$V = i_0(R_{S0} + r_0) + \tfrac{1}{2}i_{1+}r_1 \quad . \quad . \quad . \quad . \quad (59)$$

and $$0 = i_{1+}(R_{R1+} + r_0) + \tfrac{1}{2}i_0 r_1 \quad . \quad . \quad . \quad (60)$$

Therefore $$V = i_0(R_{S0} + r_0 - \frac{\frac{1}{4}r_1^2}{R_{R1+} + r_0} \quad . \quad . \quad . \quad (61)$$

and $$Z_{i0} = r_0 - \frac{\frac{1}{4}r_1^2}{R_{R1+} + r_0} \quad . \quad . \quad . \quad . \quad (62)$$

Case 8.—$R_S = R_{S0}$ at frequency ω_q but infinite at wanted sideband frequency; $R_R = R_{R1+}$ at wanted sideband frequency, but infinite at ω_q. At all other frequencies the effective impedance is zero.

The equations are simply

$$\frac{V}{R_{S0}} = v_0\left(\frac{1}{R_{S0}} + g_0\right) + \tfrac{1}{2}v_{1+}g_1 \quad . \quad . \quad . \quad (63)$$

and $$0 = v_{1+}\left(\frac{1}{R_{R1+}} + g_0\right) + \tfrac{1}{2}v_0 g_1 \quad . \quad . \quad . \quad (64)$$

Therefore $$\frac{V}{R_{S0}} = v_0\left(\frac{1}{R_{S0}} + g_0 - \frac{\frac{1}{4}g_1^2}{\frac{1}{R_{R1+}} + g_0}\right) \quad . \quad . \quad (65)$$

so that $$Y_{i0} = g_0 - \frac{\frac{1}{4}g_1^2}{\frac{1}{R_{R1+}} + g_0} \quad . \quad . \quad . \quad . \quad (66)$$

and $$Z_{i0} = \frac{1 + g_0 R_{R1+}}{g_0 + R_{R1+}(g_0^2 - \frac{1}{4}g_1^2)} \quad . \quad . \quad . \quad (67)$$

(6.2) Calculations for Table 3: Ring Modulator

Case 1.—R_S any value; R_R a constant resistance.

As explained in the main text, the input impedance of a ring modulator with constant-resistance load is a constant time-invariant resistance, and no even-order currents or voltages exist in the input loop. The equations are

$$V = i_0(R_S + r_0) + \tfrac{1}{2}I_{1+}r_1 + \tfrac{1}{2}I_{1-}r_1 + \tfrac{1}{2}I_{3+}r_1 + \ldots \quad . \quad (68)$$

and $$0 = I_{n\pm}(R_R + r_0) + \tfrac{1}{2}i_0 r_n) \quad . \quad . \quad . \quad (69)$$

giving $$Z_{i0} = \frac{V - i_0 R_S}{i_0} = r_0 - \frac{\pi^2}{16}\frac{r_1^2}{R_R + r_0} \quad . \quad . \quad (70)$$

Case 2.—R_S a constant resistance; $R_R = R_{R1+}$ at the wanted sideband frequency, but infinite at all other frequencies (strictly only at all odd-order frequencies, since no others exist in the output loop).

In this case there are even-order currents, but they all have a constant-resistance termination, R_S. Therefore eqn. (40) may be applied, namely

$$I_{1+} = \frac{\frac{1}{2}r_1[i_0(R_{S0} - R_S) - V]}{(R_{R1+} + r_0)(R_S + r_0) - \frac{\pi^2}{16}r_1^2} \quad . \quad . \quad . \quad (71)$$

Note that R_S has been given a distinctive value, R_{S0}, at frequency ω_q so that it can be shown that Z_{i0} is independent of R_{S0} although not of R_S.

Therefore $$V = i_0(R_{S0} + r_0) + \tfrac{1}{2}I_{1+}r_1 \quad . \quad . \quad . \quad . \quad (72)$$

$$= i_0(R_{S0} + r_0) + \frac{\frac{1}{4}r_1^2[i_0(R_{S0} - R_S) - V]}{(R_{R1+} + r_0)(R_S + r_0) - \frac{\pi^2}{16}r_1^2} \quad . \quad (73)$$

from which $$Z_{i0} = \frac{V - i_0 R_{S0}}{i_0} = \frac{r_0 - R_S\frac{r_1^2}{4B}}{1 + \frac{r_1^2}{4B}} \quad . \quad . \quad . \quad (74)$$

where $$B = (R_{R1+} + r_0)(R_S + r_0) - \frac{\pi^2}{16}r_1^2 \quad . \quad . \quad (75)$$

and clearly Z_{i0} is independent of R_{S0}, and depends only on the value of R_S at the even-order frequencies, $l\omega_p \pm \omega_q$, $l \neq 0$.

Case 3.—R_S a constant resistance, $R_R = R_{R1+}$ at the wanted sideband frequency, but zero at all others.

This case is exactly dual to case 2, and evidently gives

$$Y_{i0} = \frac{g_0 - \frac{1}{R_S}\frac{g_1^2}{4C}}{1 + \frac{g_1^2}{4C}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (76)$$

where $$C = \left(\frac{1}{R_{R1+}} + g_0\right)\left(\frac{1}{R_S} + g_0\right) - \frac{\pi^2}{16}g_1^2 \quad . \quad (77)$$

Case 4.—$R_S = R_{S0}$ at frequency ω_q, but infinite at all other frequencies (strictly only at all even-order frequencies); $R_R = R_{R1+}$ at the wanted sideband frequency, but infinite at all others (strictly only at all odd-order frequencies). Here the equations are

$$V = i_0(R_{S0} + r_0) + \tfrac{1}{2}I_{1+}r_1 \quad . \quad . \quad . \quad . \quad (78)$$

and $$0 = I_{1+}(R_{R1+} + r_0) + \tfrac{1}{2}i_0 r_1 \quad . \quad . \quad . \quad (79)$$

giving $$Z_{i0} = \frac{V - i_0 R_{S0}}{i_0} = r_0 - \frac{\frac{1}{4}r_1^2}{R_{R1+} + r_0} \quad . \quad . \quad (80)$$

Case 5.—As case 4, but zero replacing infinity.

Clearly this is an exactly dual case, giving

$$Y_{i0} = g_0 - \frac{\frac{1}{4}g_1^2}{(1/R_{R1+}) + g_0} \quad . \quad . \quad . \quad . \quad (81)$$

and $$Z_{i0} = \frac{1 + g_0 R_{R1+}}{g_0 + R_{R1+}(g_0^2 - \frac{1}{4}g_1^2)} \quad . \quad . \quad . \quad (82)$$

Case 6.—$R_S = R_{S0}$ at frequency ω_q, but infinite at all other frequencies; $R_R = R_{R1+}$ at the wanted sideband, but zero at all other frequencies.

There are no even-order currents in this case, so that the equations are

$$V = i_0(R_{S0} + r_0) + \tfrac{1}{2}I_{1+}r_1 + \tfrac{1}{2}I_{1-}r_1 + \tfrac{1}{2}I_{3+}r_3 + \ldots \quad (83)$$

$$0 = I_{1+}(R_{R1+} + r_0) + \tfrac{1}{2}i_0 r_1 \quad (84)$$

$$0 = I_{1-}r_0 + \tfrac{1}{2}i_0 r_1 \quad (85)$$

and for all values of n, n odd, $n \neq 1$,

$$0 = I_{n\pm}r_0 + \tfrac{1}{2}i_0 r_n \quad (86)$$

Substituting eqns. (84)–(86) into eqn. (83), we obtain

$$Z_{i0} = \frac{V - i_0R_{S0}}{i_0} = r_0 - \frac{\frac{1}{4}r_1^2}{R_{R1+} + r_0} - \frac{1}{4}\left(\frac{\pi^2}{4} - 1\right)\frac{r_1^2}{r_0} \quad (87)$$

Case 7.—$R_S = R_{S0}$ at frequency ω_q, but zero at all other frequencies; $R_R = R_{R1+}$ at the wanted sideband, but infinite at all other frequencies.

This case is dual to case 6; there are no even-order voltages, and the solution is clearly

$$Y_{i0} = g_0 - \frac{\frac{1}{4}g_1^2}{(1/R_{R1+}) + g_0} - \frac{1}{4}\left(\frac{\pi^2}{4} - 1\right)\frac{g_1^2}{g_0} \quad (88)$$

DISCUSSION ON THE ABOVE THREE PAPERS BEFORE THE ELECTRONICS AND COMMUNICATIONS SECTION, 11TH JANUARY, 1960

Mr. G. D. Monteath: Messrs. Gouriet and Newell envisage vestigial-sideband television broadcasting as the principal application for their quadrature network. In the system used now, which has been termed 'receiver attenuation', the sloping amplitude/frequency characteristic in the neighbourhood of the carrier is shaped by the receiver. The authors suggest that, given a fresh start, the alternative 'transmitter attenuation' system, in which the characteristic near the carrier is shaped by the transmitter, would be better.

The choice of receiver rather than transmitter attenuation was made in the United States on the recommendation of the National Television Systems Committee (N.T.S.C.). The N.T.S.C. proceedings* states that transmitter attenuation would have improved the signal/noise ratio by 6 dB, but this statement is

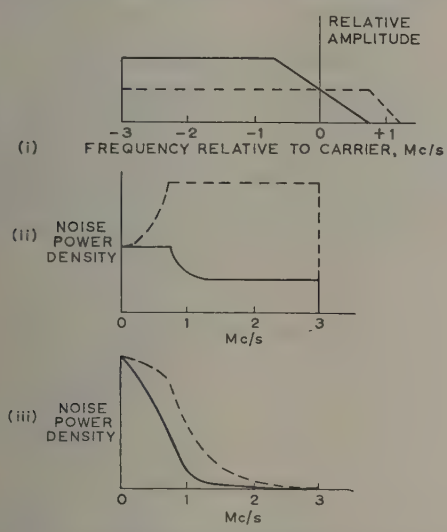


Fig. A.—Comparison between transmitter-attenuation and receiver-attenuation systems.

---- Receiver-attenuation system.
—— Transmitter-attenuation system.

(i) Transmitter amplitude/frequency characteristic.
(ii) Noise power spectrum after detection.
(iii) Noise power spectrum after detection but weighted in proportion to $(1 + f^2)^{-2}$ (f in Mc/s).

open to question. In Fig. A, (i) shows the amplitude/frequency characteristic of the transmitter for the alternative systems, according to an idealized representation of British standards. The receiver characteristics are assumed to be the same but interchanged. It can be seen from (ii), which shows the noise power spectrum after detection, that the transmitter-attenuation system gives a 6 dB reduction in noise only at the higher frequencies. There is less improvement at lower frequencies because both transmissions tend to the double-sideband type near the carrier frequency, and because a receiver designed for the transmitter-attenuation system [having a response corresponding to the broken line in (i)] admits more noise near the carrier. The total noise power for receiver attenuation is found to be only 4½ dB greater than for transmitter attenuation. Fig. A(iii) shows the effect of taking into account the reduced sensitivity of the eye to noise at higher frequencies by applying a simple weighting factor.* The advantage offered by transmitter attenuation is reduced to only 2 dB.

In fact, the N.T.S.C. recommended the receiver-attenuation system on the grounds that the transmitter-attenuation system would increase the cost of each receiver by $4–$5, requiring an additional stage of i.f. amplification and a second trap for adjacent-channel sound. Do the authors agree with this view, which argues against their proposals? It might, for example, be better to use any increase in receiver price to restore the d.c. component.

The authors propose to generate the vestigial-sideband signal by combining the two quadrature components at high power. In the first method suggested, the transmitters are connected in series, but I do not see how intermodulation could be avoided. For example, with a picture lacking fine detail, the subsidiary transmitter would have no output and its output stage would presumably be cut off. How would the output of the main transmitter pass through it? The second method, illustrated in Fig. 14, wastes half the power in a dummy load. Would it not be better to combine the two components at low power, since even linear amplification would be more efficient?

Professor Tucker and Mr. Howson are concerned with linear networks containing periodically varying resistances. Peterson and Hussey[1] have analysed these networks in terms of equivalent circuits, thereby avoiding much of the algebra. Are the assumptions made by the authors and the generality of their approach the same as in the Peterson and Hussey paper? If so, why was the equivalent circuit abandoned? Perhaps it would be too involved to be useful in the more complex cases considered.

Are second-order modulation products to be avoided only to simplify calculation, or also to improve efficiency?

Mr. B. B. Jacobsen: Paper No. 3054 E is interesting and the results are useful. The paper would have been understood by a wider circle of readers if the principle used had been explained, as it easily and accurately could have been, in terms of the response obtained from pairs of tapping points in a delay line and the combination of several such responses to synthesize an

* FINK, D. G. (editor): 'Television Standards and Practice' (McGraw-Hill, 1943).

* GILBERT, M.: 'The Visibility of Noise in Television: Part II', B.B.C. Engineering Division Monograph No. 3, October, 1955.

approximation to a square frequency response with quadrature transmission phase.

Vestigial-sideband principles appear to be used for applications requiring power, in some cases, of the order of several kilowatts and in others only milliwatts. The authors clearly are most interested in the high-power application, but in Section 7 they mention the use of vestigial-sideband signals for transmission over cables—a low-power application which is well established. For the latter purpose, the quadrature network has the advantage that the carrier frequency is not fixed by the network but can be varied from time to time if it should be required. This advantage will probably rarely be exploited but must be paid for in equipment cost. The authors' equipment appears to contain a very considerable number of components. What are the authors' views on the relative economy of the quadrature-network method and of alternative methods such as, for instance, that which uses a simple double-sideband modulator followed by a sideband shaping filter with a delay corrector?

In Section 7 the authors suggest the use of a carrier frequency of 0·4 Mc/s for transmission of a vestigial-sideband signal over a cable. This frequency is inconveniently low and it would be difficult and expensive to equalize the group delay of the cable circuit at the very low frequencies which would be required. The carrier frequency used, at least when transmitting over considerable distances, should preferably be not less than about 1 Mc/s. The C.C.I.T.T. has recommended the use of a carrier frequency of 1·056 Mc/s.* When this carrier frequency is used the lowest significant signal-frequency component on the line is of the order of 500 kc/s and it is then fairly simple to equalize the cable-circuit group delay.

In Section 1 the terms 'single sideband' and 'vestigial sideband' are used in a way that suggests that these terms are synonymous. Do the authors agree that a single sideband signal is one in which the unwanted sideband is present only as a spurious effect, but that in a vestigial-sideband signal, the vestigial sideband is present as a necessary effect and with specified attenuation? Vestigial-sideband systems can transmit signals down to and including zero frequency.

The practice of using vestigial-sideband transmission from a television transmitter would certainly reduce the demands for dispersion-free filtering made on the receiver, but it appears that the receiver would then be more open to noise at frequencies in the vestigial-sideband range. Do the authors think that the transmitted signal could be increased to compensate for this without unduly increasing the transmitter cost? Perhaps the extra cost could be justified by the smaller waveform distortion which might be expected to result when the vestigial-sideband process is carried out at the transmitter rather than at the receiver.

* C.C.I.T.T. Green Book, Vol. III bis, p. 265.

Dr. E. L. C. White: Linke's paper (Reference 9 of Pape No. 3054 E) refers to an early disclosure of the basic principl of the transversal filter in 1931. In Patent No. 517516, ther is a lengthy discourse on the use of the transversal filter, and i is rather surprising that greater use of such filters has not bee made. Linke's use is probably the best known. The schem put forward by Messrs. Gouriet and Newell is a most ingeniou application of it for generating single-sideband transmissions One other use is in producing aperture correction in televisio studio equipment.

If the taps required are symmetrical, only half the length o filter is necessary, by making use of reflection. In the single sideband case where skew symmetry is wanted, it would b necessary to have reflection at a short-circuit and taps similar t those in the left-hand half of Fig. 6. If a centre tap on th filter is required for generating the non-quadrature signal wit the appropriate delay, it is necessary to utilize the current in th short-circuit. This can be done by using the low impedance o a shunt feedback amplifier as the short-circuit.

It is doubtful whether this is the best way of making a single sideband signal. The scheme uses two modulators and th transversal filter for obtaining the quadrature component, an various other circuit-elements which are less significant. Fo about the same number of circuit-elements it is possible t employ the standard technique of modulation to give a double sideband signal clear of the original video signal, using a filte to remove one sideband. The filter must be phase linear an have a steep slope. This is another application in which on can use a transversal filter, and is described in Patent No. 51751 There is then no problem of mixing the outputs from two separat circuits and keeping the balance of gains and phases good.

Dr. D. Maurice: Referring to the paper by Messrs. Gouri and Newell, would a quadrature-signal television transmissio give rise to less distortion of waveform transients than a orthodox transmission? Calculations based on so many assump tions as these are open to argument, but I think they may we be representative. I assume that transmitter and receiver gai characteristics can be represented by maximally-flat-amplitud Butterworth functions, except for the quadrature-signal emis sion. If only minimum-phase-shift circuits are used, the grou delays corresponding to the various gain characteristics may b calculated and Table A summarizes the results.

Table A

Frequency	Transmitter		Receiver		Overall group delay		Overall gain	
	Orthodox	Quadrature	Orthodox	Quadrature	Orthodox	Quadrature	Orthodox	Quadrature
					microsec	microsec	dB	dB
$f_c - 3\frac{1}{2}$			−45*	−45*			−45	−45
$f_c - 3$	0*	0*	−3*	−3*	1·04	−1·1	−3	−3
$f_c - \frac{3}{4}$	0	−1*	−1	0	0·03	0·02	−1	−1
f_c	0*	−6*	−6*	0*	0	0	−6	−6
$f_c + \frac{3}{4}$	−2*	−20*	−19	−2*	0·13	0·15	−21	−22
$f_c + 1\frac{1}{4}$	−20*		−30	−20	−0·13	0	−50	
$f_c + 1\frac{1}{2}$	−32	−32*	−35*	−35*			−67	−67

* Specified attenuations.

From this Table, it can be seen that the overall gain charac teristics for the two systems are virtually equal. The grou delays are also nearly equal because, although in the quadratur signal case the Nyquist slope is obtained by a distortion-fre method, the contribution to overall group delay supplied by th receiver in an orthodox transmission is negligible. The tot variation of group delay, over the vision-frequency spectrum

ue to the Nyquist slope in the receiver for an orthodox transission is 0·14 microsec.

I was disappointed to find that the two methods of transmission ere so similar in this respect and I wonder if the authors have her reasons for thinking that the quadrature signal method is, fact, the better with regard to linear waveform distortion.

Professor D. G. Tucker: Messrs. Gouriet and Newell state at Levy's paper is different. The only difference seems to that he used reflections whereas the authors use taps. The flection method is more restrictive with respect to the amplide of the various components, but I cannot see any other fference in principle.

Mr. B. M. Sosin: Referring to the paper by Messrs. Gouriet nd Newell, there is one important point which has not been entioned. If the vestigial-sideband signal is produced at the igh power output of the transmitter, the limitation of the ouble-sideband transmitter as against the offset carrier transitter must be considered. The power of the transmitter epends on its boundaries. For a double-sideband transmitter iere is at least a 50% wider band, and the power is normally hree-quarters to two-thirds of that obtained with the offset arrier.

Mr. J. M. Layton (*communicated*): If, as Mr. Howson and rofessor Tucker assume, $r(t)$ is a function of carrier voltage lone, all circuits considered are linear and a cisoidal input $\varepsilon^{j\omega_q t}$ nay be assumed (as in normal circuit theory),*† instead of the uthors' $\cos \omega_q t$. All the results of the paper may then be btained in a more general form, unshackled by the assumption f resistive terminations, without in any way complicating the nalysis. Indeed, when dealing with rectangular wave variation of $r(t)$, this method, suitably developed, renders the whole of Section 6.2 unnecessary.

For instance, in the series modulator, if the source voltage is $\varepsilon^{j\omega_q t}$, then

$$i = \sum_{-\infty}^{\infty} i_m \varepsilon^{jt(\omega_q + m\omega_p)}$$

If $r'(t) \equiv r(t) - r_0$ is assumed to be any odd harmonic variation in time, so that

$$r'(t) = \sum_{-\infty}^{\infty} r'_n \varepsilon^{jn\omega_p t} \text{ and } g'(t) \equiv \frac{1}{r'(t)} = \sum_{-\infty}^{\infty} g'_n \varepsilon^{jn\omega_p t}$$

(n odd), the condition for the absence of non-zero even-order sideband components in i is found to be

$$Z_n \equiv Z[j(\omega_q + n\omega_p)] = \frac{r'_n}{g'_n}\left(Z_0 - \frac{1}{i_0}\right)^{-1} \propto \frac{r'_n}{g'_n}$$

where $Z = Z_R + Z_S + r_0$ is the total mesh impedance apart from $r'(t)$. The constant of proportionality in this relation gives $Z_0 - 1/i_0$ and, if Z_0 is known, this gives i_0. The other current components are found to be given by $i_n = g'_n(1 - Z_0 i_0)$.

In the parallel case considered by the authors, $r'(t)$ is a rectangular wave of amplitude $\frac{1}{2}(r_b - r_f)$ and therefore $r'_n/g'_n = \frac{1}{4}(r_b - r_f)^2$, so that Z_n must be constant (independent of n). In practice this means that Z_n is resistive for all n, owing to the difficulty of designing a circuit to have a constant vector impedance at all odd sideband frequencies (from $-\infty$ to $+\infty$) unless its phase is zero. This makes $Z_0 - 1/i_0$ positive real: if, however, Z_0 is complex, so also is i_0, and the supply-frequency current component is phase shifted with respect to the supply voltage.

THE AUTHORS' REPLIES TO THE ABOVE DISCUSSION

Messrs. G. G. Gouriet and G. F. Newell (*in reply*): We agree with Mr. Monteath that the improvement in signal/noise ratio would be between 2 and 3 dB. As regards the question of receiver cost we find it difficult to agree or disagree since there are so many factors involved. The point made by Mr. Monteath regarding the possibility of intermodulation being caused by connecting the two transmitter outputs in series is extremely relevant. If this could not be avoided by operating the transmitter output stages in a more linear fashion, the method would be impracticable. The difficulty in the method he suggests of generating the vestigial-sideband signal at low power and then using power amplification is, surely, to obtain such amplification without reintroducing signals in the vestigial sideband by intermodulation.

Whether harmonic synthesis, as suggested by Mr. Jacobsen, is preferable to analysis in terms of unit-impulse response is a matter of opinion. We feel that the latter approach is direct and basically simple. Furthermore it throws light on the true nature of the problem; for example, the Hilbert transform shows immediately that a knowledge of the past and future is required in order to obtain a quadrature signal. As regards the economy of components, the number of inductors and capacitors required for the production of a particular phase and attenuation characteristic is determined by the bandwidth and time delay, inherent in the characteristic, and this will be largely independent of the method used.

A carrier frequency of 0·4 Mc/s was stated merely as an example to show how the frequency spectrum of a television signal might be shifted by a small amount. It was not suggested that this frequency would be chosen for a conventional cable circuit.

We do not disagree with the definitions of 'single-sideband' and 'vestigial-sideband' as stated by Mr. Jacobsen and have used these terms not in order to suggest that they are synonymous but to stress that the treatments of both systems are identical by the method described in the paper. Single sideband is the limiting case of the vestigial sideband as regards the length of the delay line required.

The question of signal/noise ratio has been effectively dealt with by Mr. Monteath.

The method suggested by Dr. White for halving the length of delay line required necessitates the use of a loss-free line. This method prevents the possibility of compensating for loss along the line by the adjustment of the taps because the loss will not be the same at any one tap for the direct and the reflected signal. The use of a transversal filter, for the removal of the unwanted components of a double-sideband signal, must be restricted to low power because of the inherent insertion loss of such filters. The problem therefore remains of obtaining efficient, and sufficiently linear, power amplification of a vestigial-sideband signal.

We thank Dr. Maurice for his interesting contribution.

The criticism by Mr. Sosin is most important and if, in fact, it is true that the transmitter power is reduced to two-thirds as a result of increasing the bandwidth by 50%, this is sufficient to nullify any economic advantage in the method we propose.

As Professor Tucker states, the main disadvantage of the system described by Levy is the restriction imposed by the reflection method. These reflections must be kept small in order that secondary reflections are kept to reasonably small amplitude.

* Stieltjes, F. H.: 'Application of Complex Symbolism to Systems with Frequency Transposition', *Journal of the Netherland Radio Society*, 1946, **11**, p. 221.
† Zadeh, L. A.: 'Frequency Analysis of Variable Networks', *Proceedings of the Institute of Radio Engineers*, 1950, **38**, p. 291.

This aggravates the problem of subtracting the two direct components in order to obtain the relatively small quadrature component. In addition the appropriately delayed original signal is not so readily available as in our method.

Mr. D. P. Howson and **Professor D. G. Tucker** (*in reply*): In reply to Mr. Monteath, the assumptions we have made are basically the same as those made by Peterson and Hussey, and the generality of our approach is therefore the same. However, (like several other authors), we have extended the scope very greatly to consider more complex circuits, i.e. the ring modulator, and more complex terminating arrangements. The use of the equivalent circuit has no merit other than that of giving a more easily appreciated illustration of the circuit properties than is given by the equations; in all cases the equivalent circuit is deduced from the equations—and not vice versa—so that it is not very suitable as a basis for calculation. Moreover, in our cases the circuit would have to include an infinite number o branches.

There is no doubt that our main interest in avoiding even-orde modulation terms is to eliminate them from the equations i order to make solution simple, and therefore we may ofte eliminate even-order currents or voltages without necessaril eliminating both together. It is not possible to show that, i general, avoidance of even-order products leads to maximu efficiency. There are obvious practical advantages, however, i not having frequency components such as $2\omega_p - \omega_q$ to limit th available bandwidth of the $\omega_p + \omega_q$ product.

We thank Mr. Layton for his interesting contribution, an agree that the use of the cisoidal-signal representation has advan tages. Whether his method is as convenient as ours in solvin the various problems we have discussed is not at present clea We look forward to seeing his work in more detail.

1.387 : 621.316.722

The Institution of Electrical Engineers
Paper No. 3174 E
May 1960

THE CORONA DISCHARGE AND ITS APPLICATION TO VOLTAGE STABILIZATION

By E. COHEN, and R. O. JENKINS, Ph.D., F.Inst.P.

(*The paper was first received 27th July, and in revised form 12th October*, 1959.)

SUMMARY

The properties of a gas discharge which make it suitable for voltage stabilization are discussed. Various types of corona discharge are described, and it is shown that a corona discharge between a wire anode mounted axially in a cylindrical cathode is the most suitable for voltage stabilization. To obtain the required low impedance it is necessary to use hydrogen as the filling gas. The effect of dimensions on the voltage and impedance of the discharge is discussed, and characteristics of various tubes which have been developed as corona voltage stabilizers covering a range of 350–7 000 volts are shown. Finally some of the applications of such tubes and their appropriate operating conditions are given.

(1) INTRODUCTION

Gas discharge tubes have been used for a considerable time as shunt stabilizers or as the source of the reference voltage in stabilized power supply units. Tubes for this purpose usually operate with a glow discharge in an inert gas filling and often have a specially activated cathode. The potential across such a tube is generally in the range 60–150 volts and the currents through the tube in the range 1–100 mA, depending on their design. It has been realized only relatively recently apart from an early note by Medicus,[1] that the corona type of discharge can be used in a similar manner to give considerably higher stable voltages with lower operating currents. Blifford, Arnold and Friedman[2, 3] carried out an investigation on the subject in 1947, and the application of some of their tubes was described by Lichtman:[4] in this country preliminary work was done by Jacques[5] and Biram and Wade.[6] Further development on tubes operating up to 1 500 volts was described by Shelton and Wade[7] in 1953, and measurements on such tubes were made by Benson and Smith.[8] Collinson and Hill[9] carried out an investigation into the processing of similar tubes.

During the past seven years a number of tubes have been developed by the authors covering a voltage range 350–7 000 volts. It is the purpose of the paper to outline the properties of the corona discharge and to discuss the various factors which have determined the design of these tubes. The operating characteristics of some of the tubes are given, together with typical uses.

(2) GAS STABILIZER PARAMETERS

Before discussing the corona discharge in detail, the most important parameters of any type of gas discharge for voltage stabilization must be considered. The discharge tube is usually connected across the unstabilized voltage, V_I, via a series resistor R_S. The stabilized voltage, V_S, appears across the tube and feeds the external load resistor R_L, as shown in Fig. 1(*a*). The electrical equivalent of the circuit is shown in Fig. 1(*b*), where V is the voltage across the discharge tube and R is the incremental impedance of the tube, defined as dV/di, where i is the tube current. The incremental impedance is not necessarily resistive, and in the case of glow discharge tubes a fairly complex network must replace the pure resistance, R, shown in Fig. 1(*b*), if high-frequency effects are considered. In the preliminary considerations of the corona discharge, however, the impedance will be regarded as resistive, since high-frequency effects are small compared with the resistive component.

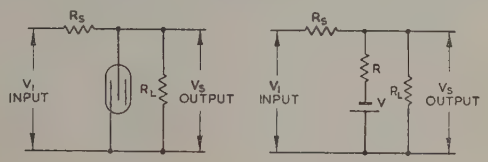


Fig. 1.—Discharge-tube circuit arrangement.
(*a*) Actual circuit.
(*b*) Equivalent circuit.

For effective operation it is desired to make the stabilization ratio, dV_I/dV_S, as high as possible, at the same time keeping the fractional loss in voltage $(V_I - V_S)/V_I$ reasonably small. It can easily be shown that

$$\frac{dV_I}{dV_S} = 1 + \frac{R_S}{R} + \frac{R_S}{R_L} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

As the voltage lost across R_S must usually be kept small compared with that across the load R_L, $R_S/R_L < 1$. Also, as $dV_I/dV_S \gg 1$, for good stabilization the smaller quantities in eqn. (1) may be neglected for most practical purposes, and hence

$$\frac{dV_I}{dV_S} \simeq \frac{R_S}{R} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

During operation with variations of input voltage and output load, the current through the tube may vary considerably, and the range of currents over which a stable discharge can exist is therefore very important.

It may thus be seen that the most important parameters of the discharge for voltage stabilization are: the discharge voltage and its stability during life, the incremental impedance dV/di, and the range of currents over which the discharge remains stable.

(3) PROPERTIES OF CORONA DISCHARGES

(3.1) Form of Discharge

A corona discharge occurs between two electrodes at low currents when there is a large electric field near one of them. In this high field, the electrons gain sufficient energy between collisions with the gas molecules to be able to ionize them on collision. The main ion formation thus takes place in a thin sheath round this electrode, and a faint glow is usually visible. These conditions can be realized by making one of the electrodes either a point or a wire, and it can be either positive or negative with respect to the other. A point electrode can pass only a

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
The paper is a communication from the Staff of the Research Laboratories of The General Electric Company, Limited, Wembley, England.

relatively small current in the corona discharge before transition to a glow or spark disharge and is thus generally unsuitable for voltage stabilization. The electrode system chosen is therefore a wire along the axis of a cylinder. To obtain the relatively high field near the wire, its diameter must be small compared with that of the cylinder. For reasons to be explained later, it is usual to make the wire the anode and the cylinder the cathode.

In the discharge, the intense ionization takes place near the anode wire within the radius S, beyond which the field is insufficient for ionization by collision to take place efficiently. Relatively few free electrons exist outside this radius, while all the positive ions formed within the sheath move out towards the cathode, and there is thus a region of positive space charge. The main voltage drop across the tube occurs within the sheath and is almost independent of the current. There is also a small voltage drop in the positive ion region due to the space charge. This latter voltage drop increases with the current and gives the true incremental impedance of the discharge. The voltage/current curve is shown diagrammatically in Fig. 2. It will be seen

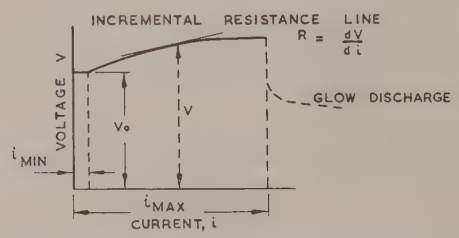


Fig. 2.—Voltage/current curve of discharge.

later that the observed value of the impedance may be rather different from this true value, owing to thermal effects. The various parameters of the discharge will now be considered in detail.

(3.2) Discharge Voltage

The breakdown between parallel plates can be treated analytically and the criterion for breakdown derived. The classical work of Townsend explained the well-known Paschen law that the breakdown voltage, V, is a function of pd, where p is the gas pressure and d the spacing between the plates. This assumes that the gas temperature is constant, and V is more strictly a function of ρd, where ρ is the gas density. This function varies with the nature of the gas and also with γ, the number of electrons released at the cathode by an arriving positive ion. An increase in γ lowers the breakdown voltage for a given value of pd. The curve of V has a minimum for a certain value of pd, which is generally known as the Paschen minimum voltage. The relationship between V and pd at values well above this minimum tends to become linear.

In the case of cylindrical electrodes a rather similar relationship can be derived.

Meek and Craggs[10] show that when the wire is the cathode there should be a theoretical relationship

$$\frac{V}{\log \frac{R}{r}} = \mathrm{f}(pr) \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where R and r are the radii of cylinder and wire.

They also show experimental results for oxygen with the wire as anode, in which $V/\log(R/r)$ is plotted against pr and a smooth curve obtained. Theoretically the varying radius of the cylindrical cathode might also be expected to have some effect, but this was evidently neglibly small. The actual form of curve is very similar to the Paschen curve for breakdo between parallel plates.

It has been found possible to plot the results we have obtain with hydrogen on this same basis, and a smooth curve (Fig. 3)

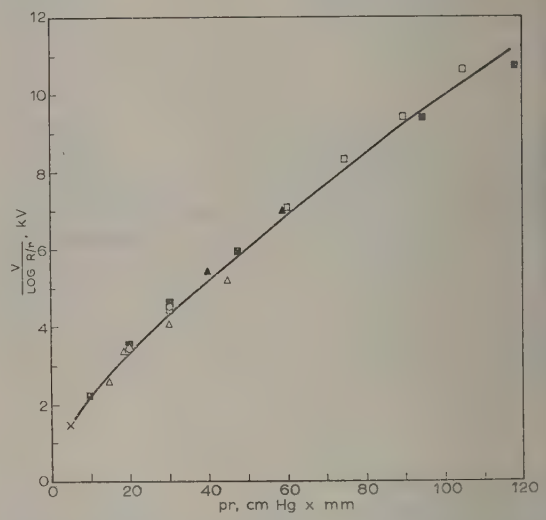


Fig. 3.—Breakdown curve for corona discharge in hydrogen.

Experimental tube dimensions, mm.

	R	r
×	15	0·5
○	3·5	0·5
▲	7	1
□	7	1·5
■	7	2·37
△	5	0·75

obtained in spite of differing values of cylinder diameter. It w be noticed that the parameters studied are above the Pasch minimum, which does not appear on the curve. One oth characteristic of the corona discharge is that its breakdo voltage exceeds its maintaining voltage by only a small amoun usually a few per cent. The values actually plotted in Fig. are the maintaining voltages at a steady current of 50 μA.

The choice of the cylinder as cathode for voltage stabilizatio is governed by the following considerations. The positive io arriving at the cathode have less energy when the cathode is cylinder than when it is a wire. This results in a lower val of γ, and the resulting discharge voltage is less dependent on and therefore less likely to change owing to changes in t cathode surface. In addition, the lower value of γ for t cylinder as cathode results in a higher voltage for a given pressu of gas, thus extending the voltage range for a given syste without exceeding atmospheric pressure. Finally, the coro discharge with the cylinder as cathode generally supports higher current before transition to a glow discharge.

The choice of gas will affect the discharge voltage for a giv pressure filling. The breakdown curve for nitrogen lies abo that for hydrogen shown in Fig. 3, that for argon slightly belo while those for helium and neon are considerably below. T choice of gas, however, is governed mainly by the need to ke a low incremental impedance, rather than by the breakdo characteristic.

(3.3) Incremental Impedance

The incremental impedance of the discharge will be inverse proportional to the length of the cylindrical system provid

at the dimensions are sufficiently accurate for the discharge to read uniformly along the wire. The value of the impedance, can be calculated approximately from the space charge outside e sheath. This calculation[11] results in the relationship

$$Z = \frac{p \log \frac{R}{r}}{2\mu Vl}\left[(R^2 - r^2) - (S^2 - r^2)\frac{\log \frac{R}{r}}{\log \frac{S}{r}}\right] \quad . \quad . \quad (4)$$

here μ is the mobility of the positive ion (its velocity at unit is pressure and in unit potential gradient), l is the system length, id S the outer radius of the ionizing sheath.

In the corona discharge $R > r$, and at low currents the sheath thin; thus $S \simeq r$. Neglecting the second-order terms in m. (4),

$$Z \simeq \frac{R^2 p \log \frac{R}{r}}{2\mu Vl} \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

o obtain an approximate idea of how Z varies with tube mensions and gas, it may be assumed that at higher voltages e breakdown relationship of eqn. (3) becomes approximately near, and thus

$$\frac{V}{\log \frac{R}{r}} \simeq bpr \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

here b is a constant for the particular gas.

Combining eqns. (5) and (6),

$$Z \simeq \frac{R^2}{2rl\mu b} \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

Remembering that R/r must be kept above a certain minimum or a corona discharge to exist, the incremental impedance of ne corona tube can be kept as low as possible by decreasing the atio R/l for the system, and by choosing a gas whose product b is large.

Mechanical considerations such as centring and rigidity of the node wire put a lower limit on R/l. With these overriding imensional limitations it is found in practice to be very desirable o choose the gas giving the largest value of μb in order to btain an adequately low impedance. Values for various gases re given in Table 1. The mobility, μ, is the ion velocity in entimetres per second in a field of 1 volt/cm at 1 cm Hg gas ressure. The parameter b is in effect a breakdown field in olts per centimetre at 1 cm Hg gas pressure. The latter data re approximate and are derived from curves of breakdown oltage in cylindrical electrode systems, given by Meek and Craggs.[10]

Table 1

Gas	Mobility μ	Breakdown parameter b	Product μb
		$\times 10^2$	$\times 10^5$
Argon ..	230	16	3·7
Helium ..	1 600	8·3	13
Hydrogen ..	1 000	30	30
Nitrogen ..	230	42	9·6
Neon ..	610	7·5	4·6

Inspection of Table 1 shows clearly that the best gas for btaining a low incremental impedance is hydrogen. As a esult the corona stabilizers which have been developed up to the present have used a hydrogen filling, apart from those for low voltages where a helium-hydrogen mixture may be preferable.

(3.4) Current Limits for the Corona Discharge

The lower limit of current for a stable corona discharge in electrode systems a few centimetres long with gas pressures less than atmospheric is usually of the order of a few microamperes. At currents lower than this the statistical fluctuations of the number of ions in the discharge are such that it can momentarily extinguish and then restrike, the tube behaving like an unquenched Geiger counter. This minimum current is in practice so low that it imposes no appreciable limitation even in the lowest power applications.

The corona discharge is unable to support more than a certain maximum current per unit length of the system, and if this current is exceeded the discharge changes discontinuously to the glow mode. The visible discharge then stretches from cathode to anode and contracts to a thin pencil while the voltage across the tube decreases. The explanation can be seen from eqn. (4) for the incremental impedance. As the current across the tube is increased more ions are necessary and the ionizing sheath becomes thicker. The second term in the square bracket of eqn. (4) can no longer be neglected and the incremental impedance therefore decreases with increasing current. When the outer edge of the sheath approaches the cathode, $S = R$ and the incremental impedance tends to zero. At this point the conditions for the transition to a glow-type discharge are reached. While this gives a qualitative explanation of the transition, it has not been possible to deduce the current at which it takes place. It can, however, be seen that in a given system the maximum current will increase with the gas pressure. The sheath thickness consists of the number of electron mean free paths necessary to form sufficient ions to support the current, neglecting those formed by photo-ionization. The mean free path is inversely proportional to gas pressure, so that at a higher pressure a greater number of mean free paths will be possible between anode and cathode before the sheath reaches out to the cathode. This will allow the maximum current in the corona discharge to be greater.

(4) DESIGN OF CORONA STABILIZER TUBES

(4.1) General Features

The foregoing theory of the corona discharge has indicated that the design of a stabilizer tube will be based on the use of a positive wire corona discharge in hydrogen. The cylindrical electrode system will be made as long as possible within an envelope of acceptable dimensions. The incremental impedance will thus be kept to a minimum and the current supported by the corona discharge to a maximum. There will be a choice of wire and cylinder radii, limited by the proviso that the wire radius must be small compared with that of the cylinder. The actual radii will be kept as small as practicable in order to keep the incremental impedance of the tube to a minimum. The lower limit on cathode radius in tubes intended for the lower voltages is imposed mainly by mechanical considerations. For higher-voltage tubes, achieving the voltage with a suitable gas pressure and dissipating the power become factors demanding a greater radius. The choice is discussed more fully in the next Section.

The usual method of assembling the tube is to employ accurately made insulating plugs which fit into the ends of the cylindrical cathode and have central holes to locate the anode rod. The plugs enclose the discharge space and prevent gas ions from being taken up by the glass envelope. The insulators are generally of ceramic, but in small tubes glass can be used without

causing serious voltage drift due to released impurities. Care is necessary to avoid unwanted discharges outside the electrode system. Some of the main features of the design can be seen from the group of tubes shown in Fig. 4. One problem in

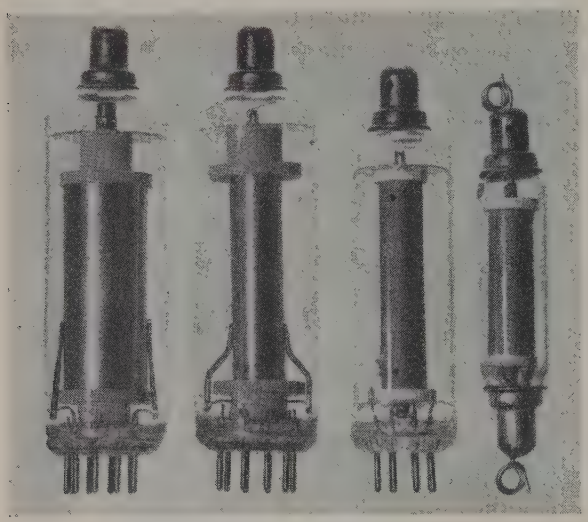


Fig. 4.—Corona-stabilizer tubes.
(*a*) Type S.C.4. (*b*) Type S.C.2. (*c*) Type S.C.1. (*d*) Type S.C.3.

processing the tubes is that their operating voltage is dependent on the pressure of hydrogen. The latter must therefore remain constant during life, and the tendency for ionized hydrogen to be absorbed by the metal parts, particularly the cathode, must be kept to a minimum. This usually means that the cathode surface treatment must be carefully controlled, and it is also usual to operate the tubes after gas filling for about 100 hours. This period is generally sufficient to stabilize the hydrogen pressure by allowing the various surfaces within the tube to achieve an equilibrium adsorbed layer.

(4.2) Choice of Dimensions

The operating voltages of tubes with different dimensions and filling pressures can be derived approximately from Fig. 3. The operating voltage is, of course, independent of the length of the system. The effect of increasing the cathode diameter with a constant anode diameter of 1 mm is shown in Fig. 5. As might be expected, the operating voltage for a given pressure increases with the diameter of the cylinder. The effect of altering the anode diameter with a fixed cathode diameter of 14 mm is seen in Fig. 6. The variation of voltage is comparatively small and there is a maximum voltage for a given pressure corresponding to an anode diameter of about 3 mm. It should be noted that the anode diameter of 4·75 mm corresponds to a cathode/anode diameter ratio of only three, which is dangerously low for maintaining a corona discharge at an appreciable current.

The effect of dimensions on the incremental impedance of the discharge at low currents can be seen approximately from eqn. (5). Values measured experimentally depend somewhat on thermal effects, which will be discussed in Section 5. The agreement between experiment and calculation using eqn. (5) is therefore not very good, the experimental values usually being about half those calculated. Experimental results from measurements made by the method of Section 5 for the three tubes with a 1 mm-diameter anode wire of similar length and three different cathode diameters are shown in Fig. 7. It will be seen that for a given

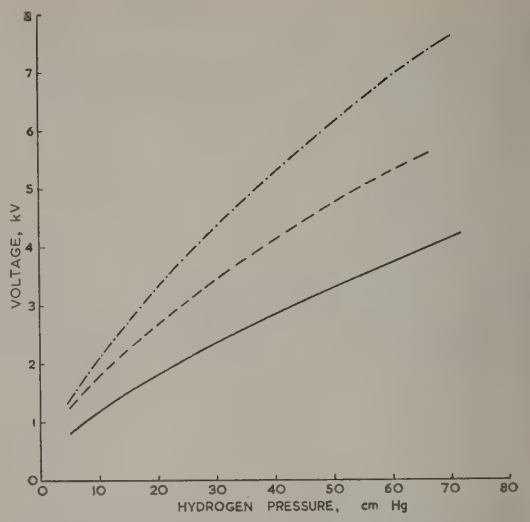


Fig. 5.—Effect of cathode diameter on discharge voltage.
Anode diameter, 1 mm. Current 50 μA.
Cathode diameter ——— 7 mm.
– – – – 14 mm.
–·–·– 30 mm.

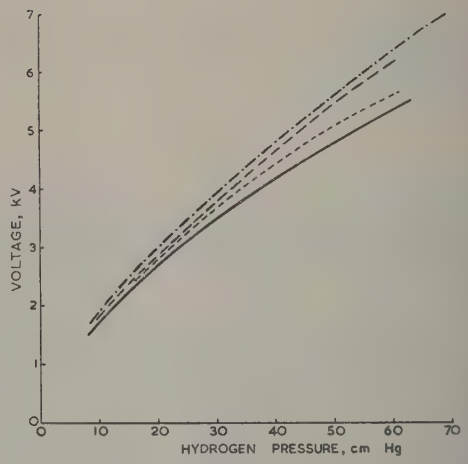


Fig. 6.—Effect of anode diameter on discharge voltage.
Cathode diameter, 14 mm. Current, 50 μA.
Anode diameter ——— 1 mm.
— — 2 mm.
–·–·– 3 mm.
– – – – 4·75 mm.

operating voltage the incremental impedance increases very rapidly with cathode diameter. The magnitude of increase is not very different from that predicted by calculation. A similar set of experimental results is shown in Fig. 8 for the variation of incremental impedance when the anode diameter is varied and the cathode diameter kept constant. This shows the desirability of keeping the anode as large as possible compatible with maintaining the corona type of discharge over an adequate current range.

The curves shown indicate that in a standard valve envelope with a type B9A base allowing a cathode diameter of about 14 mm it should be possible to make corona-discharge tubes over a voltage range of somewhat under 500 volts up to about 7 kV without exceeding atmospheric pressure for the hydrogen filling.

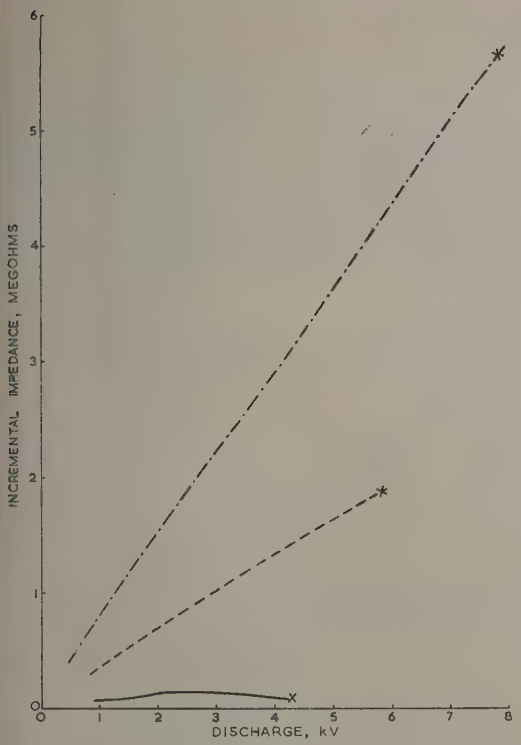


Fig. 7.—Effect of cathode diameter on incremental impedance.

Anode diameter, 1 mm.
Active length, 28 mm.
Current range, 50–200 μA.
Cathode diameter —·—·— 30 mm
– – – – – 14 mm.
——— 7 mm.
× Atmospheric pressure.

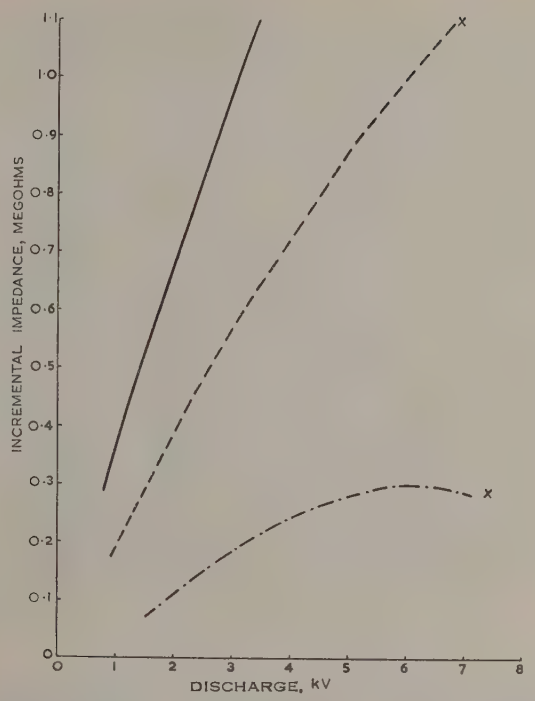


Fig. 8.—Effect of anode diameter on incremental impedance.

Cathode diameter, 14 mm.
Active electrode length, 28 mm.
Current range 50–200 μA.
Anode diameter ——— 1 mm.
– – – – – 2 mm.
—·—·— 3 mm.
× Atmospheric pressure filling.

he incremental impedances of such tubes with an active elec-ode length of about 3 cm would be of the order of 0·1 megohms, eir actual values increasing with operating voltage. The hoice of electrode diameters is a compromise depending on the nge of operating voltages to be covered by varying the gas lling pressure. As an example, from Fig. 6, a cathode diameter f about 14 mm and an anode diameter of 3 mm would be itable for voltages in the neighbourhood of 5–7 kV. This ystem, when filled at lower pressures to operate at 1–2 kV would, owever, have a higher incremental impedance than that achieved ith a smaller-diameter system filled to a correspondingly higher ressure. The latter system would also have the advantage in ractice of needing a less critical filling pressure to achieve a iven voltage tolerance. In addition, any clean up of hydrogen uring operation would have less effect on the operating voltage.

It has been found in practice that a reasonable compromise o cover the voltage range of 350–7 000 volts is to use three ifferent electrode systems. The first design (type S.C.1) has athode and anode diameters of 7 mm and 1 mm, respectively, nd covers a voltage range of 350–2 000 V. This design is ounted in a B7G based envelope. A slightly larger system ith electrode diameters of 8 mm and 1·2 mm, respectively, overs the voltage range of 2–4·5 kV. This system (type S.C.2) s mounted in a B9A based envelope. For the voltage range f 4·5–7 kV the electrode diameters are 13·2 mm and 3 mm; his system (type S.C.4) is also mounted in a B9A based envelope. n addition to these tubes a slightly shortened type S.C.1 elec-rode system (type S.C.3) is mounted in a subminiature 10 mm-diameter bulb with a flying lead at each end, and is intended mainly for portable equipment. These tubes are shown in Fig. 4.

Some of the main characteristics of such tubes will be discussed in the next Section.

A somewhat similar range of tubes, the GV5 and GV6 series, has been developed in the United States.[1] The electrode system of the GV5 tube is mounted on a B7G base and has a 6 mm-diameter cathode, while that of the GV6 tube, which covers up to 6 kV, has a B9A base. The electrical characteristics of these various tubes are rather similar to those of tubes of corresponding voltages discussed here.

(5) VOLTAGE/CURRENT AND IMPEDANCE/CURRENT CHARACTERISTICS

(5.1) Effect of Rate of Change of Current

It has been assumed so far that in the absence of clean up of hydrogen there is a unique voltage/current curve for a particular tube. This in turn implies that the incremental impedance also has a unique value for a particular current. This assumption is not strictly correct, since there are thermal effects which can be significant, particularly in the higher-voltage tubes where power dissipation becomes appreciable.

It has been seen that the voltage of any tube is a function of the gas density, particularly in the discharge sheath round the anode where most of the voltage drop occurs. If the tube is operated for a few minutes at a steady current, the temperature throughout the tube will reach equilibrium and the voltage will attain a stable value. If the current is suddenly increased the voltage will immediately rise owing to the increase in positive-

ion space charge. Subsequently, in a small fraction of a second the temperature of the gas in the discharge sheath will rise and the local gas density will decrease, causing the voltage drop in the sheath to fall slightly. As the temperature of the electrode system increases during the next few minutes of operation, gas is driven from the system to the space outside the cathode. The density of the gas in the discharge sheath thus decreases further and the operating voltage drops until a final stable value is reached. This final voltage may in some cases be below its original value at the lower current, giving an apparently negative impedance, although the true discharge impedance due to space charge is positive. As a result of these thermal effects there are three separate voltage/current and impedance/current relationships that can be derived for a particular tube. The first is for current changes taking place in less than about 0·1 sec, the second for changes in a few seconds and the third for changes maintained for a few minutes. In general, the difference between the first two relationships is relatively small but the last may differ considerably from the other two in high-voltage tubes.

The first relationship will be applicable when assessing the performance of the stabilizer tube for reducing hum, the second for reducing the effect of transients and the third for slow drifts due to variations of mains supply, circuit components or loading. The methods of measurement employed emphasize these points.

(5.2) Measurement Methods

(5.2.1) Very Rapid Changes.

The measurement for very rapid changes of current can be carried out by measuring the reduction in hum in a stabilizing circuit. The tube is operated in the circuit of Fig. 1(*a*) with a current meter in the tube lead, and hum at audio frequency (50 c/s, or greater) is fed by a capacitor on to the input point. The ratio of hum at the input and output can be measured by an oscillograph or valve voltmeter, and the tube impedance, R, can be calculated from eqn. (1), R_L being the impedance of the measuring device which is the output load in this case. Measurements can be carried out over the current range, and the relationship between R and the current i obtained. The voltage/current curve can be derived by the integration

$$V_i = \int_{i_0}^{i} R di + V_0$$

where V_0 is the voltage for the lowest stable current i_0. As, however, fluctuations in current due to hum are in general small, the values of R are usually adequate for assessing the tube performance without needing the full voltage/current curve.

(5.2.2) Short Transient Changes.

One method of deriving the voltage/current curve for transient changes is to operate the tube in the circuit of Fig. 1(*a*) with a voltage divider and potentiometer across the output and a current meter in the tube lead. The tube current is then set at the minimum stable value, i_0, and the voltage V_0 across the tube measured. The current is then increased to a new value, i, and the corresponding voltage, V_i, measured as quickly as possible, i.e. within a few seconds. The current is then returned to i_0 and the voltage allowed to return to its original value, V_0. The procedure is repeated for various values of current so that the full curve can be plotted. The incremental-impedance/current curve can be derived by differentiation.

(5.2.3) Slow Changes.

The voltage/current curve is derived by measuring V with the potentiometer, allowing the tube to operate at each current for a few minutes until the voltage is stable. The incremental-impedance relationship can, in this case also, be derived differentiation.

(5.3) Typical Characteristics

The S.C.1 system is filled to various pressures to give a ran of tubes operating at voltages of 350 volts up to 2 kV. The voltages are usually measured at a current of 250 μA and t individual tube voltages are kept to within $\pm 2\frac{1}{2}\%$ of th nominal value, or to within ± 20 volts for tubes below 800 vol

Typical voltage/current and impedance/current curves f 400- and 2000-volt tubes are shown in Figs. 9 and 10. In t

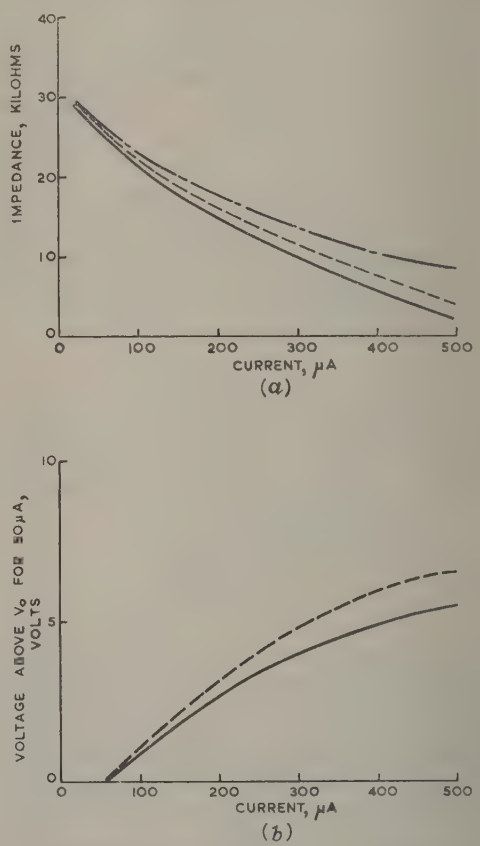


Fig. 9.—Curves for 400-volt S.C.1 tube.
(*a*) Impedance/current curves.
(*b*) Voltage/current curves.
---- Fast potentiometer changes.
—— Slow potentiometer changes.
—·— 100 c/s a.c. changes.

former curves, the voltage for 50 μA is taken as the zero. will be seen that for the 400-volt tube there is little differen between the three impedance/current curves derived by t different methods. For the higher-voltage tubes the curves f the slow changes lie considerably below those for the fast changes.

The voltage/current relationships for the various voltage tub for transient changes are shown together in Fig. 11. The ve low position of the 350- and 400-volt curves relative to t 600-volt one results from the inclusion of a small quantity helium, which enables the voltage tolerances to be achiev more easily. The curves for the subminiature tubes are ve similar to their miniature counterparts, but owing to the short electrode system, the impedances are about 30% higher.

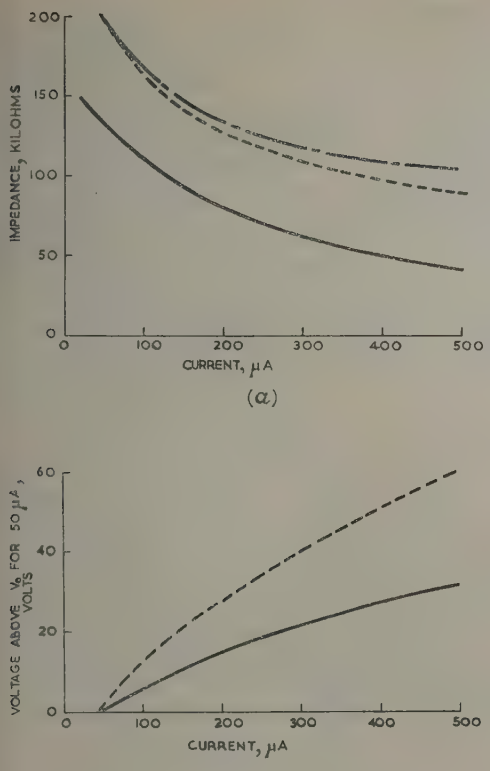


Fig. 10.—Curves for 2 kV S.C.1 tube.
(*a*) Impedance/current curves.
(*b*) Voltage/current curves.
– – – – Fast potentiometer changes.
——— Slow potentiometer changes.
—·— 100 c/s a.c. changes.

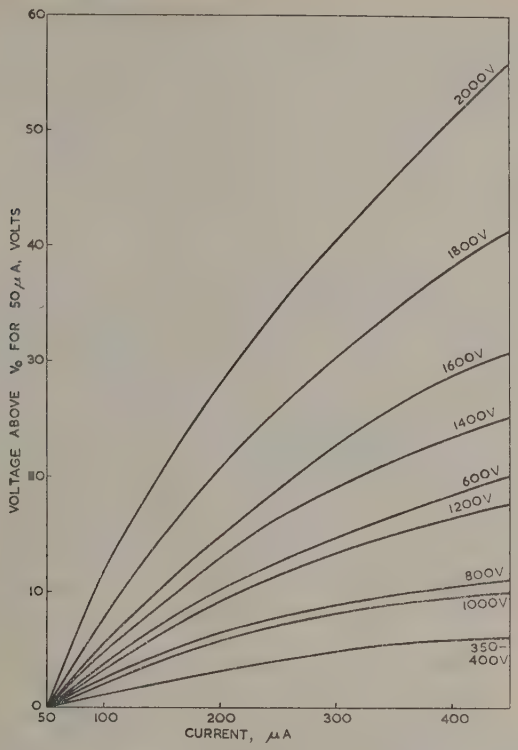


Fig. 11.—Voltage/current curves for S.C.1 tubes with various voltages.

The S.C.2 system is usually filled to give operating voltages f 2·5, 3·0, 3·5 and 4 kV with ±2½% tolerance. Typical mpedance/current and voltage/current curves for the 4 kV tubes re shown in Fig. 12. It will be seen that the difference between he curves has increased, and in the case of the 4 kV tube, a egative impedance is observed for slow changes below about 00 μA.

The S.C.4 system is usually filled to give operating voltages f 5, 6 and 7 kV with a 2½% tolerance. Typical impedance/cur-ent and voltage/current curves for a 7 kV tube are shown in ig. 13. This latter tube has a negative impedance for slow hanges above about 1·5 mA.

(6) MINIMUM CURRENTS

As outlined previously there is a minimum current below hich the corona discharge becomes unstable. This instability best observed by means of a cathode-ray oscillograph. The inimum currents below which instability sets in are shown in able 2. As the minimum current is proportional to the ystem length, the figures for the S.C.3 system are lower than hose for the S.C.1 with corresponding voltages.

(7) MAXIMUM CURRENTS

As previously explained, there is also a maximum current hat the corona discharge can pass, above which the transition to the glow discharge takes place. To observe the change it is

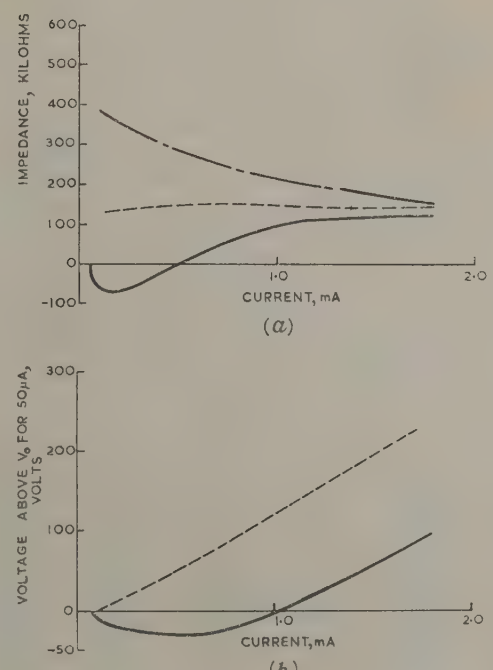


Fig. 12.—Curves for 4 kV S.C.2 tube.
(*a*) Impedance/current curves.
(*b*) Voltage/current curves.
– – – – Fast potentiometer changes.
——— Slow potentiometer changes.
—·— 100 c/s a.c. changes.

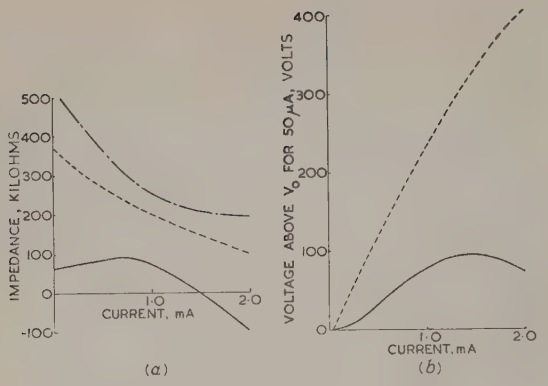


Fig. 13.—Curves for 7 kV S.C.4 tube.

(*a*) Impedance/current curves.
(*b*) Voltage/current curves.
---- Fast potentiometer changes.
—— Slow potentiometer changes.
—·— 100 c/s a.c. changes.

Table 2

Minimum Current for a Stable Discharge

Nominal voltage	Current	
	Type S.C.1	Type S.C.3
volts	μA	μA
350	1	1
400	1	1
600	4	3
800	11	7
1 000	14	9
1 200	16	11
1 400	16	11
1 600	16	11
1 800	16	11
2 000	16	11

Minimum current for the S.C.2 and S.C.4 systems is 25 μA independent of voltage, within their usual ranges.

usually sufficient to observe either the tube voltage or current as the input voltage in the circuit of Fig. 1(*a*) is slowly increased. A sudden drop in tube voltage or rise in tube current can be seen at the transition. The discontinuity is very marked in higher-voltage tubes but is barely apparent in the lowest-voltage ones, the apparent impedance just becoming increasingly negative. In this case an oscillograph connected across the tube shows that the discharge has changed, a sawtooth trace due to a relaxation oscillation between the two discharge states appearing.

In the high-voltage tubes the maximum current that the corona discharge will support decreases if the tube is operated near the maximum for any length of time. This effect appears to be due to the increase in electrode temperature, and if the tube is allowed to cool, the original maximum value is recovered. This means that, although a particular high-voltage tube might be able to pass momentarily a corona discharge current of 2 mA when first switched on, this figure would be considerably reduced if the tube was operated continuously much above about 1 mA.

One other effect worth noting is that operation at high currents will tend to increase the rate of clean up of hydrogen during the life of the tube. This effect is noticeable mainly in the lower-voltage tubes where the rate of change of voltage with pressure is higher. It becomes more serious in the subminiature tube because of the smaller volume of hydrogen, and for this reason a maximum continuous operating current well below the upper limit for the corona discharge is recommended. In practice this is not a serious drawback, as the subminiature tubes are generally operated in small portable units where the available current is in any event low.

The various maximum current ratings for different tubes are given in Table 3. It may be seen from the figures for the S.C.1 system that the peak current that the corona discharge will pass increases with the voltage rating. This was to be expected from the discussion in Section 3, since the filling pressure increases with voltage and the ionizing sheath thickness for a given current decreases. Beyond a certain pressure, however, the maximum current reaches a value almost independent of pressure, and this may be seen for the higher-voltage tubes in Table 3. The reduction in the continuous rating compared with the peak maximum is more marked in the higher-voltage tubes owing to their higher electrode temperature.

Table 3

Maximum Current

Type S.C.1			Type S.C.3			Type S.C.2			Type S.C.4		
Nominal voltage	Maximum current		Nominal voltage	Maximum current		Nominal voltage	Maximum current		Nominal voltage	Maximum current	
	Peak*	Continuous†		Peak*	Continuous‡		Peak*	Continuous§		Peak*	Continuous§
volts	μA	μA	volts	μA	μA	volts	μA	μA	volts	μA	μA
350	400	400	350	350	100	2 500	1 750	1 000			
400	500	500	400	400	100				5 000	2 000	1 000
600	600	600	600	450	100						
800	600	600	800	450	100	3 000	1 750	1 000			
1 000	650	650	1 000	500	100				6 000	2 000	1 000
1 200	700	700	1 200	600	100						
1 400	750	750	1 400	650	100	3 500	1 750	1 000			
1 600	800	800	1 600	650	100				7 000	2 000	1 000
1 800	850	850	1 800	650	100						
2 000	900	900	2 000	650	100	4 000	1 750	1 000			

* Peak should never be exceeded as a change of discharge mode may occur.
† May be operated for prolonged periods at currents approaching their peak values without the peak values becoming lower.
‡ Maximum recommended continuous current is to ensure an adequate life; operation above this current may result in excessive drift of operating voltage.
§ Maximum continuous ratings should not be exceeded for any appreciable time, as this results in a serious lowering of the peak current.

(8) EFFECT OF AMBIENT TEMPERATURE

The ambient temperature has a relatively small effect on the operating voltage of corona tubes. This might be expected since the total quantity of the gas in the tube does not alter, and its mean density must therefore remain constant with temperature. It can be seen qualitatively, however, that if the temperature of the bulb is raised the temperature of the gas near it will tend to rise slightly more than that of the gas within the electrode system. This will result in the density of the gas outside the system decreasing and that inside increasing, as the pressure throughout the inside of the envelope is constant. This small increase in gas density within the system causes the voltage to rise, and it will thus have a small positive temperature coefficient. In practice, this coefficient is usually between 0·01 and 0·02% per deg C in the range −20 to +70° C and is fairly constant with temperature.

(9) NOISE

All types of gas discharge produce noise at lower frequencies owing to the random fluctuations associated with the ionizing process, in addition to noise at higher frequencies associated with the high electron temperature in the discharge plasma. The random fluctuations of ionization naturally become more significant as the discharge current and the number of ions associated with it decrease. The fluctuations result in random fluctuations of the voltage across the discharge. These constitute white noise, which increases with decreasing current.

Measurements have been made using an oscillograph with a 2 Mc/s bandwidth, estimating the r.m.s. noise voltage from the visual display of 'grass' on a millisecond time base. Table 4

Table 4

NOISE

Type	Nominal voltage	Noise output measured at 200 μA
	volts	MV (r.m.s.)
S.C.1	350	6
	400	7
	600	12
	800	30
	1 000	41
	1 200	70
	1 400	84
	1 600	86
	1 800	85
	2 000	76
S.C.2	2·5 kV	60
	3·0 kV	60
	3·5 kV	60
	4·0 kV	60
S.C.4	5·0 kV	125
	6·0 kV	150
	7·0 kV	165
S.C.3	Noise from the S.C.3 tube at 50 μA will be approximately 1·5 times that for the corresponding S.C.1 tubes	

shows these values for various tubes for a current of 200 μA. It was noted that the r.m.s. noise voltage increased rapidly to about three times the values shown as the current approached its lower stable limit. As the current increased above 200 μA the noise decreased slowly to about two-thirds of its value when the current approached its maximum.

One effect also noticed was a coherent oscillation usually at a frequency of about 1–2 Mc/s occurring at high currents in some of the higher-voltage tubes. The r.m.s. amplitude of these oscillations was of the order of 1 volt. These outputs are due to oscillation of the positive-ion plasma as a whole. The precise conditions which allow this to happen are not known, but the observed frequency agrees quite well with that calculated from the estimated ion density.

(10) STABILITY DURING LIFE

As indicated previously, stability of the voltage during life is entirely a function of the hydrogen pressure, assuming that the initial hydrogen is pure. To ensure this, spectroscopically pure hydrogen is generally used. Any changes in voltage during operation will generally be in a downward direction due to clean up of hydrogen. These changes are likely to be less for higher-voltage tubes, owing to the decrease in the rate of change of voltage with pressure at higher pressures. Changes will also be less the greater the volume of the envelope and the smaller the operating current.

A typical life-test curve for a tube operating at 1 kV and 250 μA is shown in Fig. 14. The life of a tube may be usefully

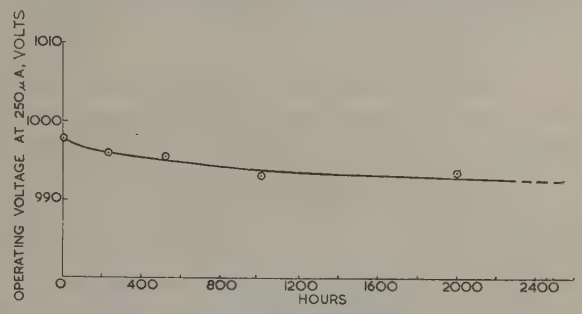


Fig. 14.—Curve of voltage during life for a typical 1 kV tube.

defined as the time taken for the voltage to alter by the tolerance originally allowed. Taking this as a basis, lives of most tubes operating at 250 μA should be in excess of 10 000 hours. The lower-voltage tubes may drift outside this tolerance in about 5 000 hours, but this useful life would increase at lower currents. The subminiature tubes operating with currents of about 100 μA will have life performances similar to their type B7G counterparts operating at 250 μA. The high filling pressures and larger volumes of the tubes with type B9A envelopes should ensure a life of 10 000 hours at currents of about 500 μA.

(11) APPLICATIONS OF CORONA TUBES

From the foregoing data it will be seen that one of the main applications for corona tubes is as shunt stabilizers for high-voltage power units supplying currents under 1 mA. For the tubes with voltages under 2 kV the series resistance is likely to be of the order of 1 megohm for most applications. This value will give a voltage stabilizing ratio of better than 10 for input fluctuations, and the tube itself will give the supply a source impedance of less than 100 kilohms, the exact value depending on the rate of fluctuation of the load. The tubes with voltages above about 2 kV will generally require a somewhat higher series resistance to achieve a worthwhile stabilization ratio, and a value of 2–5 megohms is likely to be suitable.

The small over-voltage necessary to strike the discharge with negligible delay is nearly always considerably less than the voltage dropped by the series resistor when current is passing. The

inclusion of a small amount of thoria on the insulator[12] has proved effective in eliminating time delays in striking when low-voltage tubes are used in very-low-current circuits.

Typical applications are as shunt stabilizers in the power units for Geiger–Müller tubes, photomultipliers, cathode-ray tubes and other low-current electron-beam devices. Tubes have also been used for other specialized purposes such as stabilizing the auxiliary discharge in a radar t.r. switch tube and for protecting capacitors from an over-voltage during the initial warming cycle.

Apart from their use as shunt stabilizers they can also be used as voltage-reference tubes. In this case, the operating current will best be of the order of 200 μA except for the subminiature tube where it should be between 50 and 100 μA.

A fortuitous property of the corona discharge is that the impedance and maximum currents both impose approximately the same practical limits to the use of the tubes. An illustration of this point is that with a 1 kV tube, the 1-megohm series resistor necessary to give a useful stabilizing ratio drops 650 volts at maximum current. This is about the maximum that most users would be prepared to tolerate in order to achieve stability, and so a much larger maximum current would be unusable. A rather similar relationship exists between the parameters in tubes at other voltages.

(12) CONCLUSION

The paper has outlined the main characteristics of corona discharges and has shown how they can be employed for voltage stabilization. Tubes with conventional valve envelopes can be made to stabilize voltages in the range 350–7 000 volts. Such tubes can pass between half and one milliampere.

In conclusion the possibilities of extending these limits will be considered briefly. The upper voltage limit can be extended simply, by operating the tubes in series. The voltage of a single tube, filled at less than atmospheric pressure, could also be increased by increasing the electrode diameters. An alternative but better method, in that the tube would be smaller and have a lower impedance, would be to use a gas filling above atmospheric pressure. This, however, would necessitate a more expensive method of tube construction.

It is not possible to increase the current by operating tubes in parallel, as this would involve extremely accurate matching of the operating voltages and very small over-voltages to strike the discharge. In a single tube a longer electrode system will, however, increase the maximum current and decrease the impedance correspondingly. A more compact tube with a h current rating can be made by using a helical anode wire mou coaxially within a larger-diameter cylindrical cathode, the charge sheath round the anode thus having the required leng As an example, a tube has been made with a cathode of le 48 mm and diameter 29 mm. The anode is a 20 mm-diar helix with six turns of 1·2 mm-diameter wire. This operated at about 5 kV when filled with hydrogen to a pre of 50 cm. It was able to pass approximately twice the co current possible in a conventional tube with the same cat length, filling pressure and voltage, and also had a some lower impedance.

It may be possible by some of these methods discusse increase the number of applications for which corona stab tubes can be used.

(13) ACKNOWLEDGMENT

The authors wish to thank the M.O. Valve Co., Ltd., Claudgen, Ltd., for permission to publish the paper.

(14) REFERENCES

(1) MEDICUS, G.: *Zeitschrift für Technische Physik*, 1933 p. 304.
(2) BLIFFORD, I. H., ARNOLD, R. G., and FRIEDMAN, H.: U States Naval Research Laboratories Report No. 1947.
(3) BLIFFORD, I. H., ARNOLD, R. E., and FRIEDMAN, *Electronics*, 1949, **22**, p. 110.
(4) LICHTMAN, S. W.: *Proceedings of the Institute of E Engineers*, 1951, **39**, p. 419.
(5) JACQUES, T. A.: A.E.R.E. Memo EL/M/42.
(6) BIRAM, J., and WADE, F.: A.E.R.E. Memo EL/M/52.
(7) SHELTON, E. E., and WADE, F.: *Electronic Engineering*, 1 **25**, p. 18.
(8) BENSON, F. A., and SMITH, J. P.: *Journal of Scie Instruments*, 1953, **30**, p. 192.
(9) COLLINSON, A. J. L., and HILL, D. W.: *ibid.*, 1955, **32**, p
(10) MEEK, J. M., and CRAGGS, J. D.: 'Electrical Breakdow Gases' (Clarendon Press, 1953), Chapter 2.
(11) LOEB, L. B.: 'Fundamental Processes of Electrical Disch in Gases' (Wiley, 1939), Chapter 10C.
(12) JENKINS, R. O.: British Patent Application No. 20864,
(13) COHEN, E., and JENKINS, R. O.: British Patent Applic No. 30093, 1958.

621.372.81 : 621.385.632.3

The Institution of Electrical Engineers
Paper No. 3229 E
May 1960

A SIMPLE METHOD FOR PREDICTING THE CHARACTERISTICS OF TAPE STRUCTURES

By J. ALLISON, Ph.D., B.Sc.(Eng.), Graduate.

(*The paper was first received* 3rd *June, and in revised form* 3rd *November*, 1959.)

SUMMARY

A method is described for calculating the dispersion characteristics and coupling impedance of a Karp-type slow-wave structure. The structure is regarded as a ridge waveguide, periodically loaded with short-circuited stub lines. The analysis is simple and lends itself to the speedy assessment of the performance of a particular circuit.

Measurements on 8 and 4 mm experimental backward-wave oscillators are described and are shown to verify the usefulness of the approximate theory. These results are also compared with those obtained by field-theory analyses.

(1) LIST OF PRINCIPAL SYMBOLS

l = Tape width.
$2l'$ = Slot width.
w = Ridge width.
$D = 2l$ = Pitch.
d = Ridge-to-tape distance.
β_f = Phase-change coefficient of fundamental space harmonic.
β_n = Phase-change coefficient of nth space harmonic.
θ = Phase change per section of composite structure.
ϕ = Phase change per section of unloaded structure.
λ_0 = Wavelength in free space.
λ_g' = Wavelength in ridge waveguide.
λ_c = Cut-off wavelength.
X_s = Lumped loading reactance per section.
Z_g = Characteristic impedance of unloaded waveguide.
Z_0 = Characteristic impedance of stub line.
Z_{00} = Characteristic impedance of equivalent-T network.
C = Travelling-wave-tube gain parameter.
Z_c = Coupling impedance.
E_0 = Field strength at the centre of a gap.
I_0 = Beam current.
V_0 = Beam voltage.

(2) INTRODUCTION

The most successful slow-wave circuit used for the construction of millimetre-wave backward-wave oscillators is the so-called Karp structure[1] (Fig. 1), which is a ridged waveguide with

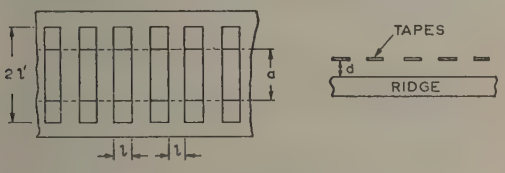


Fig. 1.—Slow-wave structure.

transversely slotted wall opposite the ridge. This is because of the relative ease of construction of the circuit. In predicting the performance of a valve using such a slow-wave structure, it is essential to know the dispersion characteristics of the structure. These are most conveniently shown by plotting frequency against phase shift from one slot to the next. This is the dispersion curve, or ω/β diagram, for the structure. The slope of the radius vector at any point on the dispersion curve, ω/θ, is then proportional to the phase velocity, v_{ph}, and the slope of the curve at any point, $d\omega/d\theta$, is proportional to the group velocity, v_g. The phase-change coefficient for any space harmonic can be calculated from the usual formula:

$$\beta_n D = \beta_f D \pm 2\pi n \quad . \quad . \quad . \quad . \quad . \quad (1)$$

or

$$\beta_n = \beta_f \pm \frac{2\pi n}{D}$$

where n is any integer.

Thus it is only necessary to know the shape of the dispersion curve for the fundamental, since the performance in any other space-harmonic régime can be deduced from this. For instance, the first backward wave has a phase shift per section equal to 2π minus the phase shift of the fundamental (i.e. $n = -1$).

(3) METHODS OF OBTAINING DISPERSION CURVES

Since the structures used in millimetre-wave valves are very small and the field decays rapidly away from the tapes, it is not possible to measure their dispersion characteristics directly. A very reliable method of obtaining the dispersion curve is to make measurements on a scale model. With this model, which can be several times bigger than that actually used, it is much easier to use probe techniques to measure the variation of field strength along the structure. The phase-change coefficient, β_f, at a given frequency, f, can then be deduced by measuring the wavelength on the structure. This can be done directly using a probe or a perturbation technique. The dispersion curve of the final structure can be obtained very simply by direct linear scaling from that of the model.

Whilst the scale-model method is useful for evaluating the performance of a finalized structure, it is rather cumbersome when used for the initial design. Consequently, a method for calculating the shape of the ω/β diagram, even if only approximately, is very desirable.

Butcher[2] has shown that this can be done by field-theory analysis of an idealized structure. The analysis is, however, complex and does not lend itself to estimating the effect of changing any of the structure dimensions. Pierce[3] has carried out an approximate field-theory analysis by replacing the tapes by a sheet conducting in the transverse direction only.

Karp[1] considers the structure as approximating to a ladder line of shunt capacitance and series stub lines. The shunt capacitance is that between the ridge and one of the tapes, and the gap between any two tapes is equivalent to a series stub. It is very simple to calculate the phase change per section, and hence β_f for such a transmission-line model.

A closer approximation can be obtained by regarding the structure as a ridge waveguide, periodically loaded with stub lines. The analysis is simple and provides an easy method for determining the change in shape of the dispersion curves with varying circuit parameters. Although this approach is not so rigorous

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
Dr. Allison is in the Department of Electrical Engineering, University of Sheffield.

as the field theory one, it is much simpler when used to assess the performance of a new circuit design.

(3.1) Calculation of Dispersion Characteristics

The slow-wave structure can be considered as a ridged waveguide, periodically loaded with short-circuited stub transmission lines. The well-known formula for the phase shift in a periodically loaded waveguide[4] is then applicable:

$$\cos\theta = \cos\phi - \frac{1}{2}\frac{X_s}{Z_g}\sin\phi \quad . \quad . \quad . \quad . \quad (2)$$

The model which will be considered is an equivalent ridge waveguide, of length D per section, with a characteristic impedance Z_{eq} and loaded with a reactance X_s per section. Since the capacitance and inductance of the actual length per section, $D/2$, have been distributed over an equivalent length, D, it follows that

$$Z_{eq} = 2Z_g \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The guide wavelength and characteristic impedance of the ridge waveguide are given by

$$\lambda_g' = \frac{\lambda_0}{\sqrt{\left[1 - \left(\frac{\lambda_0}{\lambda_c}\right)^2\right]}} \quad . \quad . \quad . \quad . \quad (4)$$

and

$$Z_g = Z_{0\infty}\frac{\lambda_g'}{\lambda_0} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where $Z_{0\infty}$ is the characteristic impedance for the H_{01} mode at infinite frequency.

Also, the lumped reactance of the stub line is given by

$$X_s = \frac{Z_0'}{2}\tan\frac{2\pi l'}{\lambda_0} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

Eqns. (3)–(6) can now be substituted in eqn. (2) to give the phase change per section:

$$\cos\theta = \cos\frac{2\pi}{\lambda_g'}D - \frac{Z_0'}{8Z_g}\tan\frac{2\pi l'}{\lambda_0}\sin\frac{2\pi D}{\lambda_g'} \quad . \quad . \quad (7)$$

The wavelength in the ridge waveguide at cut-off can be calculated by assuming the cross-section to be equivalent to an infinitely-wide composite parallel-strip transmission line with a discontinuity capacitance at the junction and calculating the transverse resonance of such a line. Curves of cut-off frequency and impedance have been derived in this manner by Cohn.[5] These enable values of λ_c and $Z_{0\infty}$ to be obtained directly if the waveguide dimensions are known.

(3.2) Calculation of the Characteristic Impedance of the Stub Line

The usefulness of the foregoing analysis is dependent on the evaluation of the characteristic impedance of the series stub line. This can be estimated by considering the tapes to be infinitely thin. Let x be the direction in the plane of the tapes, parallel to their narrow dimension, let y be perpendicular to the plane of the tapes and take the origin at the gap centre. Let us now consider the complex harmonic function:

$$\text{arc sin}\,(x + jy) = u + jv$$

where u is the potential function and v is the flux function. This function describes the entire orthogonal field geometry between two semi-infinite coplanar conducting planes, separated by a gap from $x = -1$, to $+1$ (Fig. 2). If we assume that the potential and charge at $x = +2$ are very little changed if the

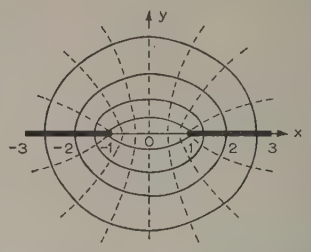


Fig. 2.—Geometry of the fields between successive tapes.

conducting planes are terminated at $x = \pm 3$, the function ca also describe the field between two parallel coplanar strips, th width of each being equal to the gap width.

In the plane of the tapes, $y = 0$, and $u = \text{arc sin}\,x$ an $v = \text{arc cosh}\,x$.

These equations lead to the characteristic impedance of th stub line Z_0' given by:

$$Z_0' = \sqrt{\frac{\mu_0}{\epsilon_0}}\frac{\pi}{\text{arc cosh}\,2} = 450 \text{ ohms}$$

(3.3) Calculation of the Coupling Impedance

Consider one section of the tape structure of Fig. 1. If th field is assumed to decrease as the square root of distance awa from the edge of a tape the quasi-static field in the gap can the be written[7]

$$E(x) = E_0\Big/\left[1 - \left(\frac{2x}{l}\right)^{1/2}\right]^{1/2} \quad . \quad . \quad . \quad . \quad (8$$

where E_0 is the field strength at the centre of the gap ($x = 0$) The space harmonic associated with the field, which has a periodicity of $2l$, can be found by a Fourier expansion at a specific time, as follows:

$$\exp(j\beta_0 x)E(x) = \sum_{-\infty}^{+\infty} C_n \exp\left(-j\frac{2\pi n}{2l}x\right)$$

or

$$E(x) = \sum_{-\infty}^{+\infty} C_n \exp(-j\beta_n x) \quad . \quad . \quad . \quad . \quad (9$$

By the usual integration, the coefficients are shown to be

$$C_n = \frac{E_0\pi}{4}J_0\left(\frac{\beta_n l}{2}\right) \quad . \quad . \quad . \quad . \quad . \quad (10$$

It is now convenient to replace the ridge waveguide model with the equivalent filter model of Fig. 3. The equivalent-T repre-

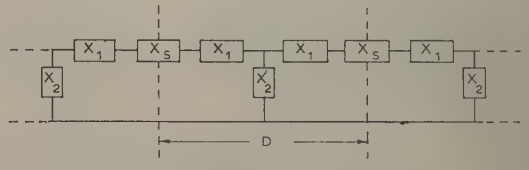


Fig. 3.—Filter analogue of slow-wave circuit.

sentation of the ridge waveguide has series and shunt element given by

$$X_1 = jZ_0\tan\frac{\pi}{\lambda_g}2l$$

$$X_2 = Z_0\Big/j\sin\frac{2\pi}{\lambda_g}2l$$

e power into the filter, which is assumed to be correctly ter-nated, is

$$P = i^2 Z_{00}$$

ere Z_{00} is the characteristic impedance of the network. The ltage across any gap is then

$$|V \text{ gap}| = \sqrt{\left(\frac{P}{Z_{00}}\right)} X_s = \int_{l/2}^{l/2} \frac{E_0 dx}{\sqrt{\left[1 - \left(\frac{2x^2}{l}\right)\right]}} \quad . \quad (11)$$

The value of E_0 obtained for this equation can be substituted eqns. (9) and (10) to give

$$E(x) = \frac{X_s}{2l}\sqrt{\left(\frac{P}{Z_{00}}\right)} \sum_{-\infty}^{+\infty} J_0\left(\frac{\beta_n l}{2}\right) \exp(-j\beta_n x) \quad . \quad (12)$$

The magnitude of the first backward-wave harmonic can be tained from eqn. (12) by putting $n = -1$:

$$|E_{-1}| = \sqrt{\left(\frac{P}{Z_{00}}\right)} \frac{X_s}{2l} J_0\left(\frac{\beta_{-1} l}{2}\right) \quad . \quad . \quad . \quad (13)$$

The circuit impedance for this space harmonic is then given by

$$Z_{c(-1)} = \frac{|E_{-1}|^2}{2P(\beta_{-1})^2} = \frac{\left[X_s J_0\left(\frac{\theta_{-1}}{4}\right)\right]^2}{2Z_{00}(\theta_{-1})^2} \quad . \quad . \quad (14)$$

Since the dispersion curve can be obtained as outlined pre-ously, the dependence of θ_{-1} on frequency is known, and the rcuit impedance at any frequency can be calculated from qn. (14). The gain parameter, C, of the tube can then be btained from

$$C = Z_{c(-1)} \frac{I_0}{4V_0} \quad . \quad . \quad . \quad . \quad . \quad (15)$$

lso, if the cold loss of the circuit is known and the magnitude f the space harmonic is calculated from eqn. (13), the current equired to start oscillations for a given active length of structure an be calculated.[1]

(4) EXPERIMENTAL BACKWARD-WAVE OSCILLATORS

Two main types of experimental backward-wave oscillator have een used to ascertain the usefulness of the approximate calcula-ions. Both incorporate Karp-type structures. Type 1 oscil-ates in the 8 mm band and type 2 in the 4 mm band. The neasured dimensions of each structure are given in Table 1.

Table 1

	Type 1	Type 2
	in	in
Guide width	0·159	0·125
Slot width, $2l'$	0·159	0·064
Ridge-tape gap, d	0·016	0·015
Ridge width, w	0·070	0·062
Tape width, l	0·009	0·004
Pitch, $D = 2l$	0·018	0·008
Guide depth	0·080	0·061

The 8 mm circuit is fabricated by winding and brazing molyb-enum tape round a molybdenum block in which the appro-riate waveguide dimensions have been milled. The 4 mm tructure is made by photo-etching the ladder-line slots into thin molybdenum sheet, this being subsequently brazed to a milled molybdenum block. This structure is designed so that the overall width and depth of the ridge waveguide are identical with those of rectangular waveguide 26. This enables the transition from the ridge waveguide to waveguide 26 to be reduced to a simple ridge taper.

The initial valves have been made demountable because of ease of adjustment of the structure and to allow simple cathode recoating. The experimental arrangement is shown in Fig. 9.

The totally immersed guns used for the 8 mm and 4 mm valves are designed to provide beams of 0·050 × 0·025 in and 0·060 × 0·010 in, respectively. Both guns use oxide-coated cathodes, and the grid structure is so arranged as to allow access to the cathode surface for recoating. Beam focusing is achieved by immersing the gun in an axial magnetic field produced by an electromagnet generating a flux density of up to 3000 gauss, with a pole separation of 6 in.

Standard Q-band cartridge-type crystals are used to detect the 8 mm oscillations. These are included in a waveguide mount and tuned by an adjustable short-circuiting backing plunger. Before experimental work of any kind could be started in the 4 mm band, suitable detectors for these frequencies had to be developed. The design eventually chosen is one in which a silicon crystal and tungsten cat's-whisker are positioned across waveguide 26 by an adjustable mount, tuning again being by a variable backing plunger.

The 8 mm oscillator has been c.w. operated throughout the experiments, but initial experiments with the 4 mm valve have been conducted using a pulsed anode voltage owing to its estimated high starting current. Voltage pulses of up to 60 microsec at a repetition frequency of 3 kc/s are applied to the valve. This repetition rate has been chosen so that a selective amplifier, tuned to the same frequency, can be used in conjunction with the crystal detector.

(5) COMPARISON OF EXPERIMENTAL AND CALCULATED RESULTS

(5.1) General Performance

A very convenient way of assessing the performance of a backward-wave oscillator is by obtaining the output spectrum in the form of an oscillogram. This is achieved by applying an accelerating direct voltage with a 50 c/s alternating voltage superimposed. The oscillograph is adjusted to display the signal detected by the crystal during the interval when the voltage sweeps over the tuning range of the valve. In this way it is possible to vary one of the operating parameters of the valve and instantly to observe the effect on the whole of the spectrum. Oscillograms obtained in this manner for the 8 mm valve are shown in Fig. 4. These are seen to contain fine-line structure which is caused by multiple reflections on the slow-wave circuit and in the output waveguide, both many wavelengths long. These effects are exaggerated, since no attempt has been made at this stage to match the circuit to the output waveguide, other than tapering the ridge waveguide to the dimensions of the output waveguide. Since the reflections in the output waveguide are small and passive, the resulting contribution to the fine structure can be reduced by inserting attenuation, but this is at the expense of maximum power output.

(5.2) Frequency/Voltage Spectrum and Dispersion Curves

Since cavity wavemeters, calibrated over the entire 8 and 4 mm bands covered by the two oscillators, are not at present available, the oscillation frequencies have, in most cases, been found from measurements of guide wavelength, using appropriate voltage-standing-wave indicators.

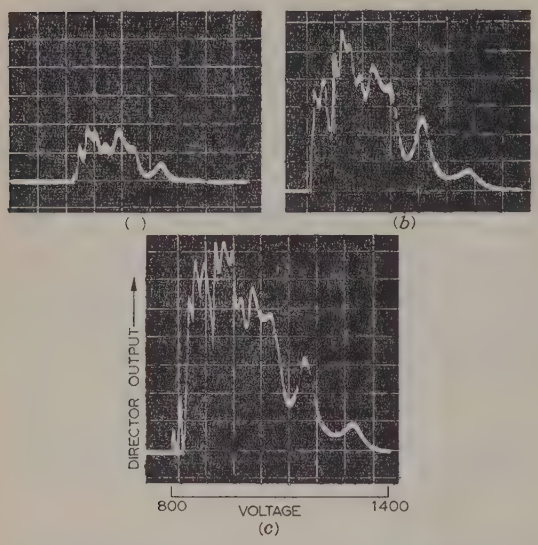


Fig. 4.—Effect of increase in beam current on bandwidth and detector output for 8 mm backward-wave oscillator.

(*a*) 0·7 mA (approx.).
(*b*) 1·2 mA (approx.).
(*c*) 1·5 mA (approx.).

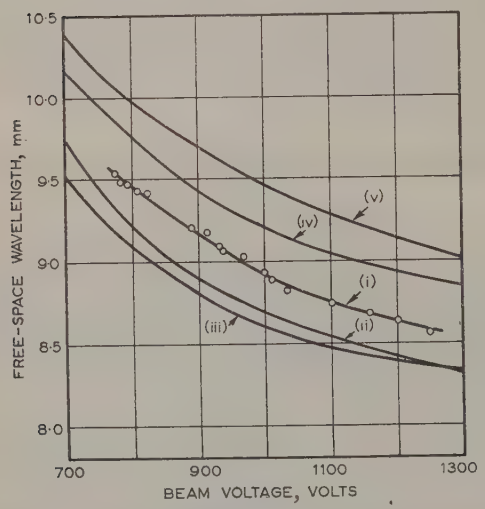


Fig. 5A.—Free-space wavelength versus beam voltage for 8 mm backward-wave oscillator.

(i) Experimental results.
(ii) Field-theory calculation (Butcher).
(iii) Field-theory calculation (Pierce).
(iv) Ridge-waveguide approximation.
(v) Lumped-capacitance approximation (Karp).

(5.2.1) 8 mm Oscillator.

The measured frequency/voltage spectrum for the 8 mm oscillator is shown in Fig. 5A. The oscillation bandwidth of over 12% requires accelerating voltages of 800–1 400 volts. This bandwidth is limited at the high-frequency end by a rapid increase in circuit loss, and at the low-frequency end by the falling off of circuit impedance.

The dispersion curve, calculated by the method outlined previously, is compared with that calculated by field-theory analyses and that obtained experimentally from voltage/frequency measurements in Fig. 6A. The waveguide dimensions have

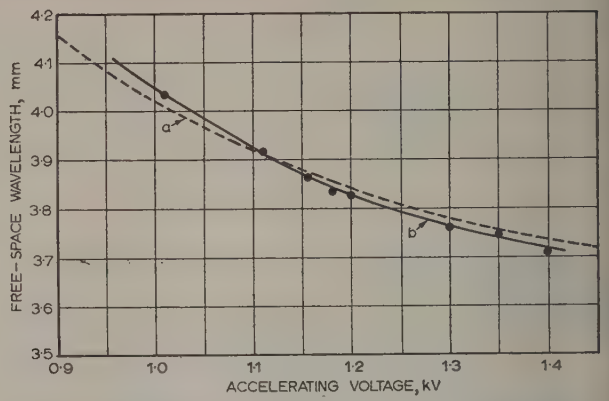


Fig. 5B.—Free-space wavelength versus beam voltage for 4 mm backward-wave oscillator.

(*a*) Calculated using ridge-waveguide approximation.
(*b*) Experimental results.

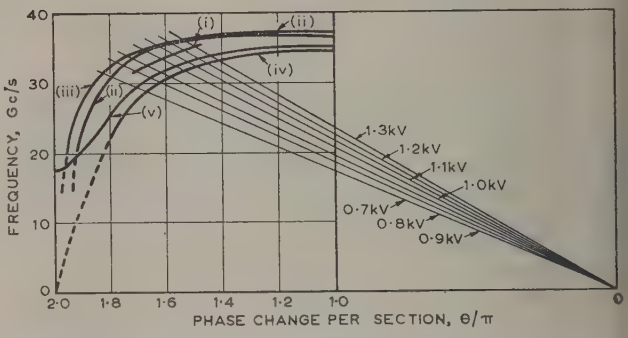


Fig. 6A.—Comparison of theoretical and measured dispersion curves of 8 mm valve.

(i) Experimental results.
(ii) Field-theory calculation (Butcher).
(iii) Field-theory calculation (Pierce).
(iv) Lumped-capacitance approximation (Karp).
(v) Ridge-waveguide approximation.

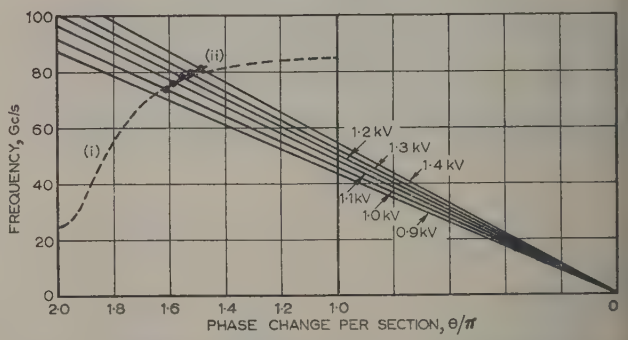


Fig. 6B.—Comparison of theoretical and measured dispersion curves for 4 mm valve.

(i) Calculated using ridge-waveguide approximation.
(ii) Experimental results.

been measured to an accuracy of within $\pm 0{\cdot}001$ cm, thus making any errors in the calculation due to this cause quite insignificant. We see that there is quite good agreement between the approximately calculated curve and the experimental one. Further, it seems that there is little to choose between the two methods of calculation as regards agreement with the experimental results.

The theoretical frequency/voltage curves, as obtained from the

lated dispersion curves, have been plotted in Fig. 5A for parison with that obtained experimentally. Here, again, approximate theory is in quite good agreement with the rimental results.

) 4mm Oscillator.

e measured tuning curve for the 4mm oscillator is shown ig. 5B. It will be noted that oscillations occur over at a 10% band of frequencies. The dispersion curve for the cture, obtained from these results, is compared in Fig. 6B that calculated using the previous theory. The resulting g curve is included in Fig. 5B for comparison with the rimental results. The excellent agreement between theory experiment is perhaps a little fortuitous, since approximations e in the theory could possibly lead to larger discrepancies.

(5.3) Power Output

1) 8mm Oscillator.

he power output of the 8mm oscillator has been measured g an 8mm thermistor in a standard Q-band mount, with ciated tuning piston. The change in resistance of the mistor when absorbing microwave radiation is measured by conventional Wheatstone network, which also lends itself irect calibration of the thermistor, using d.c. power absorp-. It will be appreciated that thermistors are not the most rate means of measuring Q-band power. However, the ticular bridge used has been calibrated against a radiation-ssure wattmeter, and it is therefore thought that measurements le with the thermistor are sufficiently accurate to give a good ication of the power level.

he power output of the valve is particularly dependent on strength and location of the focusing field and on the beam rent. If these parameters are kept constant, a typical power-put/wavelength curve is of the form shown in Fig. 7. How-

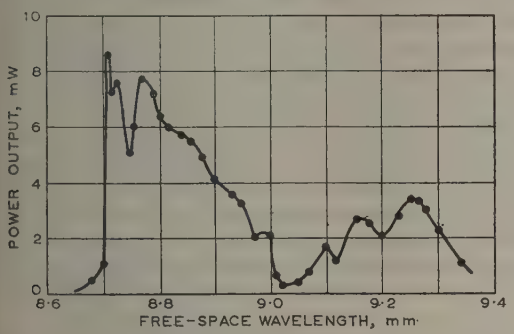


Fig. 7.—8mm backward-wave-oscillator power output versus wavelength.

Detector and magnet adjusted to maximize measured power output at $\lambda_0 = 9{\cdot}25$ mm.

er, if the variables are adjusted at each voltage so as to maxi-ze the measured output, a characteristic as shown in Fig. 8 is tained. This shows that an output greater than 10mW over 5% band and greater than 2mW over a 10% band are quite ssible, whilst the maximum output is around 18mW.

An approximate formula for estimating the power output of a ckward-wave oscillator is[1]

$$P_{out} = 86{\cdot}5J_0^{3/2}V_0^{5/4}(W\lambda_0^2) \text{ milliwatts} \quad . \quad . \quad (16)$$

ere J_0 is measured in amperes/cm^2, V_0 is in kilovolts and all lengths are in centimetres ($W/2$ = width of electron beam). e estimated power output of the tested valve, as found by

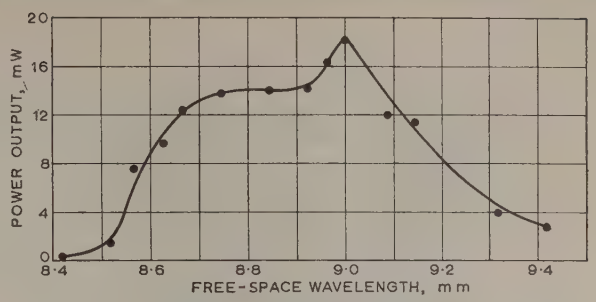


Fig. 8.—8mm backward-wave oscillator power output versus wavelength.

Detector and magnet adjusted to maximize measured power output at each frequency.

inserting the measured current density, beam voltage and wavelength in the above formula are given in Table 2.

Table 2

λ_0 mm	P_{out} mW
9·4	18
9·1	24
8·9	29
8·7	35

Comparing these with the measured values shown in Fig. 8, it is seen that the estimated power output is larger than the measured power output at the centre of the band. This is due, in some part, to ambiguity in the choice of beam width used in the calculation.

(5.3.2) 4mm Oscillator.

Because of the pulsed operation of the 4mm oscillator, its mean power output is very small and has been of insufficient magnitude to be recorded on a thermistor-bridge wattmeter. Also, the only thermistor mount available is designed for operation at 8mm wavelengths and nothing is known of its sensitivity to 4mm radiation. However, the power output of the valve can be estimated using eqn. (16). Table 3 shows the result of this computation.

Table 3

λ_0 mm	P_{out} mW
4·0	20
3·8	32
3·7	42

From these results, and from the experience gained from the measurements at 8mm, it is thought that the power output of the 4mm oscillator will not exceed 20mW.

(5.4) Starting Currents

(5.4.1) 8mm Oscillator.

It has been found difficult to measure the beam current of the valve accurately because of interception of the beam by the tapes. The collector current is thus not a reliable indication of the current in the beam; in fact, in most cases, all of the beam current is intercepted before reaching the collector. However, the starting current can be estimated by increasing the negative grid voltage until the oscillations just disappear and noting the collector current both at this time and when the beam is deflected by the focusing magnet, so as not to intercept the structure. In

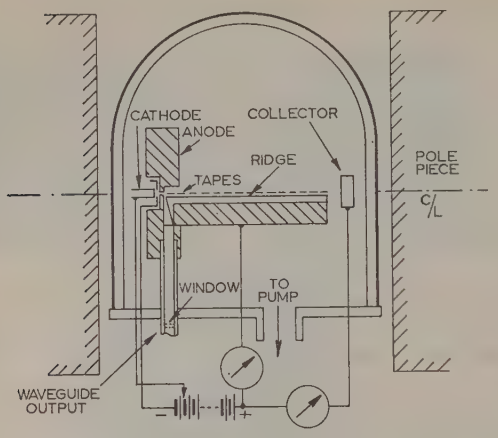


Fig. 9.—Experimental arrangement of 8 and 4 mm oscillators.

this way, it has been found that the starting current over the frequency band is of the order of 1 mA.

If we consider the particular case of an accelerating voltage of 1 kV, we find that, for this particular tube,

$$C_{start} \simeq 0{\cdot}02$$

The active lengths of circuit required to maintain oscillations may be calculated using this value of C, provided that the cold circuit loss is known.[1] For the particular structure studied the dependence of the active length on the cold loss is illustrated in Table 4.

Table 4

Attenuation/beam wavelength	L
dB	in
0	0·31
0·3	0·43
1·0	0·59

(5.4.2) 4 mm Oscillator.

Starting current densities for the 4 mm valve are given in Table 5.

Table 5

λ_0	J_{start}
mm	amp/cm²
4·0	2·8
3·9	2·8
3·8	3·5
3·7	3·9

From these results, the gain parameter, C_{start}, with an anod voltage of 1 kV is found to be 0·02. Using this figure, th active length of the circuit, as shown in Table 6, is again foun to be less than 1 in for all but the very high cold-circuit-loss case

Table 6

Attenuation/beam wavelength	L
dB	in
0	0·15
0·3	0·20
1·0	0·25
2·0	1·00

(6) CONCLUSIONS

It has been shown that the representation of a Karp-typ structure by an equivalent ridge waveguide periodically loade with short-circuited stubs provides a very convenient way o predicting the performance of this particular structure. Whils it is not claimed that the technique is as fundamental as the field theory approach, it is much simpler and quicker, and give comparable results.

(7) ACKNOWLEDGMENTS

The author is indebted to Prof. A. L. Cullen for his constan help and encouragement. Thanks are also extended to th Royal Radar Establishment for providing some of the 8 mm slow-wave structures and, in particular, to Dr. H. Duckwort for his helpful advice. Acknowledgments are also due t Ferranti, Ltd., and to the General Electric Co., Ltd., for supplyin cathode components.

The work described has been carried out under Admiralt Contract C.P. 12191/57.

(8) REFERENCES

(1) KARP, A.: 'Backward Wave Oscillator Experiments a 100–200 kMc/s', *Proceedings of the Institute of Radi Engineers*, 1957, **45**, p. 496.

(2) BUTCHER, P. N.: 'A Theoretical Study of Propagation alon Tape Ladder Lines', *Proceedings I.E.E.*, Paper No. 2294 R March, 1957 (**104** B, p. 169).

(3) PIERCE, J. R.: 'Propagation in Linear Arrays of Paralle Wires', *Transactions of the Institute of Radio Engineers* 1955, **ED–2**, p. 13.

(4) BRILLOUIN, L.: 'Wave Propagation in Periodic Structures (Dover Publications, 1946).

(5) COHN, S. B.: 'Properties of Ridge Waveguide', *Proceedings of the Institute of Radio Engineers*, 1947, **35**, p. 783.

(6) WEBER, E.: 'Electromagnetic Fields' (Wiley, 1950).

(7) MOTZ, H.: 'Electromagnetic Problems of Microwave Theory (Methuen, 1951).

..317.61: 621.382.3

The Institution of Electrical Engineers
Paper No. 3206 M
May 1960

MEASUREMENT OF TRANSISTOR CHARACTERISTIC FREQUENCIES IN THE 20–1 000 Mc/s RANGE

By J. BICKLEY, B.Sc.(Eng.), Graduate.

(*The paper was first received 28th August, and in revised form 6th November*, 1959.)

SUMMARY

Apparatus is described for the rapid determination of the cut-off equencies, f_1 and f_α, of transistors in the 20–1 000 Mc/s range. ccurate measurements at these frequencies are made possible by the plication of transmission-line techniques to the method of com- ring the high-frequency voltages which appear across small resistors nnected to two leads of the transistor. Methods are adopted which parate the measuring circuits from the input circuit and make the sign of the latter non-critical. The relative accuracy of f_1 and f_α easurements is discussed, and it is concluded that the inherently ore accurate f_1 measurement should have an error within $\pm 5\%$, hereas the error in f_α is probably 2–3 times higher. A few typical easurements are given.

(1) INTRODUCTION

As part of a programme of research into very-high-frequency licon transistors, apparatus was required which would provide rapid assessment of their upper frequency limit. The para- eters most commonly used for this purpose are f_α, the frequency which the modulus of the emitter-to-collector short-circuit rrent gain, α, falls to $1/\sqrt{2}$ of its low-frequency value, and f_1, ie frequency at which the modulus of the base-to-collector rrent gain, β, becomes unity.

Commercial apparatus is available which will measure 4-ter- inal admittance parameters over the frequency range required, ut it is slow in use and very costly. Several equipments ave been described in the literature,[1–5, 8] but most do not reach sufficiently high frequency, and the methods used seem to be ear their ultimate frequency limit. It was therefore decided at the basic method of injecting current into the transistor and mparing the voltage drops across small measuring resistors in ie appropriate branches would be adapted to wholly trans- ission-line techniques. The resulting apparatus has been used ccessfully up to 1 Gc/s, which is the frequency limit of the ceiver and signal generator available rather than that of the easuring jig itself.

(2) DESIGN PRINCIPLES

In order to reduce stray immittances to a minimum, coaxial ansmission lines were used to connect the measuring resistors points on the transistor leads as close as possible to the header. was decided that a more satisfactory design would result from jecting current into one lead of the transistor and measuring the other two, rather than attempting to inject and measure in ne lead. This would enable the two measuring limbs to be ade closely symmetrical; the f_1 measurement would then be ade with the collector and base leads in series with the easuring arms, the signal current being injected into the nitter, and the f_α measurement made with the emitter and llector leads in the measuring arms, using base injection.[6]

The measuring resistances should be as low as possible, but since it was desirable to use standard commercial components wherever possible, the characteristic impedance of the system was chosen as 72 ohms; since the measuring resistors are in parallel with the measuring circuit, their effective value becomes 36 ohms. Although this is higher than achieved elsewhere,[3] in practice it is not a serious disadvantage, because low-power transistors usually have a very low collector capacitance rather than a very low base resistance.

In previous circuits it has been necessary to provide constant-current injection to overcome the change in the transistor input impedance when switching the low-impedance voltmeter from one measuring resistor to the other. At the frequencies contemplated it is very difficult to preserve sufficient series impedance using a resistor,[4] and the transformer injection method[3] involves either low signal/noise ratio or the use of a powerful generator—which, in turn, is liable to give rise to leakage currents. These difficulties were avoided by inserting identical 20 dB attenuator pads in the output from each measuring resistor. As the 72-ohm input resistance of the receiver is switched from the outer end of the pad, there is a change of only 1% in the impedance of the measuring circuit, and the change in input impedance of the transistor will be still less. Thus it is unnecessary to provide a constant-current source.

(3) CONSTRUCTION

The circuit adopted is shown in Fig. 1, and in its physical form (Fig. 2) consists of a terminal block containing three tapered converging 36-ohm coaxial lines which emerge close together at the top of the block. The centre conductors are hollow and form the sockets for transistor leads of normal length, so providing the smoothest transition feasible from the coaxial line to the transistor header. The header for which it is designed can be brought within 2 mm of the block, to which are plugged and bolted the measuring arms 2 and 3 and the input arm 1, the joints being designed for minimum discontinuity.

The three arms are designed to be identical and each consists of a short 36-ohm line containing a blocking capacitor, leading to a T-junction each end of which bears a standard 75-ohm type-N socket. On each arm, the measuring resistor (mounted in a type-N plug) is connected to one side of the T-junction and the attenuator pad to the other. Identical *RC* decoupling networks are connected to the inner conductors of the three arms above the blocking capacitors through holes in the outer conductors. This enables the more stable common-base biasing to be used whether f_1 or f_α is being measured, and helps to maintain symmetry. For convenience, each arm is split longitudinally. Before final assembly, admittance measurements were made on the three arms, and the two most similar were used as the measuring arms. The remaining arm is used for the input circuit, after removal of the unwanted socket.

The outputs of the two attenuators on the measuring arms are fed to a communication receiver via a relay-operated coaxial selector switch. This enables the layout of the apparatus to be arranged without regard to accessibility.

Written contributions on papers published without being read at meetings are vited for consideration with a view to publication.
Mr. Bickley is at the Services Electronics Research Laboratory.

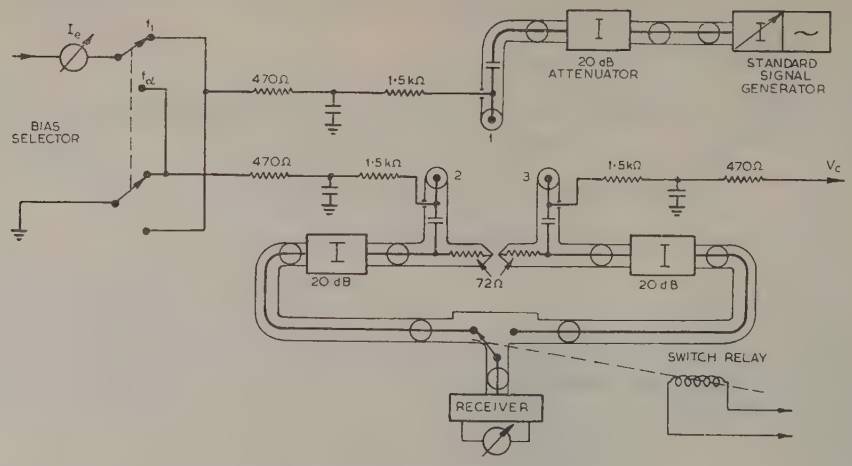


Fig. 1.—Current-gain measurement circuit.

Transistor lead	Socket number	
	f_1	f_α
Emitter	1	2
Base	2	1
Collector	3	3

All capacitors are 2200 pF.

Two receivers and one signal generator are used to cover the 20–300 Mc/s range. At high frequencies the narrow pass-band of the receiver and the frequency fluctuations of the system make tuning very difficult, and so another generator and a panoramic receiver are used in the 300–1 000 Mc/s range.

The main requirements for the generator are an accurate attenuator calibrated at close intervals and accurate frequency calibration.

(4) TESTS ON THE APPARATUS

Since the method involves a comparison of the signals received from two channels, any differences between the channels may cause errors. These differences may occur in the measuring arms and in the attenuators and switch.

Admittance measurements on the measuring arms showed that there is in each arm a shunt capacitance of about 0·2 pF associated with the biasing network, and that the series inductance of the blocking capacitor is negligible at 240 Mc/s. Bearing in mind that the equality of the two arms is more important than their absolute value, it is considered that errors due to these causes should be negligible up to at least 1 Gc/s.

The accuracy of the 20 dB fixed attenuators is quoted by the manufacturers as ±0·5 dB up to 1 Gc/s, with a voltage standing-wave ratio better than 1·1 up to the same frequency; once again equality is required rather than accuracy. The coaxial switch does not have a constant characteristic impedance, but this will not produce severe errors so long as the electrical path lengths from the two attenuators to the receiver are equal and reasonably short.

The difference in the two measuring channels was measured as a function of frequency by applying equal input signals to each channel and comparing the outputs. This test was carried out by short-circuiting together the three leads of a transistor header and inserting this into the coaxial-line sockets in the usual way. Measurements over the range 20–1 000 Mc/s showed that the differences of 0·2 dB at 20 Mc/s fell gradually to zero at 130 Mc/s and remained zero up to 360 Mc/s. It was then found to rise from 0·1 dB at 400 Mc/s to 0·5 dB at 500 Mc/s

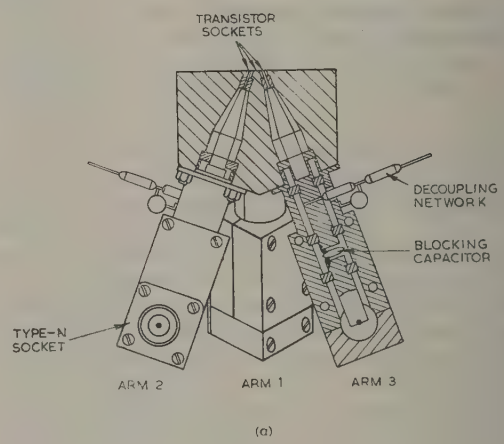


Fig. 2.—Sectioned views of jig.
(*a*) Showing the measuring arms 2 and 3 and (behind) the input arm 1.
(*b*) Side view.

to 0·9 dB at 1 Gc/s, with minor irregular fluctuations of 2 dB or less. The reproducibility after small changes in the ion of the short-circuiting header was ±0·1 dB.

rther tests were made to investigate the effects of header fringing capacitances. Two of the wires of a transistor er were short-circuited so as to link the input socket to of the measuring arms. The remaining open-circuited wire adjusted so as to imitate its position when carrying a istor. The ratio of the signals in the two measuring arms ound to increase with frequency at the expected 6 dB/octave, the capacitance was calculated to be 0·38 pF, corresponding ratio of −28 dB at 500 Mc/s.

ie input and measuring-arm sockets were next linked by the test possible length of wire. The signal ratio in this case uated with frequency and averaged −60 dB. When the ets were not linked, the coupling was estimated at −80 dB. as concluded that the direct coupling between the sockets far outweighed by that due to the transistor header.

(5) OPERATION AND EXPERIMENTAL ACCURACY

fter the transistor is inserted in the coaxial-line sockets the ing conditions are adjusted. The emitter direct current is layed continuously, but the collector voltage must be read robes across the sockets, owing to the voltage drop in the upling resistors. To avoid subsequent change, stabilized er supplies are used, and there is no difficulty in maintaining stancy to within 0·1 volt for several hours.

he signal generator and receiver are then tuned, with the erator giving an output of less than 5 mV. The receiver ut is adjusted to a reference value when receiving the lower he two inputs from the measuring arms. After switching he other input the signal generator output is attenuated l the reference point, modified to include the zero error, is in reached. The difference of attenuator readings gives $|\alpha|$ $\beta|$ directly if the transistor voltage distortion is neglected. a maximum signal level of 5 mV the error due to voltage ortion will not exceed 0·4 dB at the largest voltage ratios, with a 3 dB voltage ratio given by the attenuator the current o will be 3·11 dB. This corresponds to an error of $+2\frac{1}{2}\%$, which can be corrected.

ince the transistor input impedance would be a poor match he 50-ohm signal generator, a 20 dB 75-ohm attenuator pad sed to connect the 75-ohm coaxial cable from the generator he input of the jig. Standing waves are thus confined to the rt length of line between the pad and the transistor header, so do not affect the calibration of the attenuator.

age and Boothroyd[9] have shown that α may be deduced n measurements of $|\alpha|$ and $|\beta|$, and it has also been shown[7] for most transistors a knowledge of f_α, f_1 and α_0 (the low-uency value of α) is sufficient to predict the variation of the nitude and phase of α with frequency. Both f_1 and f_α can measured with the apparatus described, but for a rapid parison of different specimens of one design there are many ntages in measuring f_1 rather than f_α. These are:

(*a*) Depending on the impurity distribution in the base, f_α may 20–100% higher than f_1, which adds to the difficulties of easurement.

(*b*) At f_1 the two signals being compared are equal; there can erefore be no errors due to amplitude distortion or attenuator accuracy.

(*c*) The rate of change of log $|\beta|$ with log frequency near f_1 is out twice as great as that of log $|\alpha|$ near f_α, which gives a higher curacy for the f_1 measurement for a given attenuator or zero ror.

(*d*) The slope of the log $|\beta|$/log-frequency curve is constant over wide range of frequencies for most types of transistor. Thus an timate of f_1 may be made when it lies outside the frequency range of the apparatus, or is obscured by the feedback due to the parasitic elements of base spreading resistance (together with the measuring resistance) and collector capacitance, which may cause a rise in the measured $|\alpha|$ or $|\beta|$ at high frequencies. This property may also be used to smooth out random errors in the experimental results.

It is a common expedient to extend the useful range of transistor measuring apparatus by introducing zero corrections when deviations from the ideal performance occur. This is permissible, provided that the corrections are reproducible and, in particular, that they are not dependent on the properties of the transistor under test. If this condition is not met, the corrections are only approximate and must not be large.

In most previously described equipments the errors arise in the parts of the circuit closely linked to the transistor, so the corrections must be kept small. However, in the apparatus described here the errors associated with the measuring arms are negligible up to high frequencies and the deviations found occur in parts of the circuit which are effectively shielded from the transistor. It is therefore considered that larger zero corrections are permissible with this apparatus than with others and that the upper frequency limit is determined, not by the absolute value of the correction, but by its rate of change with frequency, and by the reproducibility.

In view of these factors it is considered that the error in the measurement of external f_1 with this apparatus is less than ±5%, but the error in external f_α may be higher by a factor of two or three, owing to possible errors in the attenuation measurements.

The errors arising from the measuring resistances R_1 and R_2 and the header capacitance C_H can be estimated by postulating an idealized transistor having no internal parasitic impedances. In this case the correction to the measured value, f_m, of f_1 is approximately [8]

$$2\pi f_m C_H(R_1 + R_2) \times 100\%$$

If $C_H = 0{\cdot}2$ pF the correction will be 10% when f_m is 1·1 Gc/s. This figure could be improved by adopting a lower characteristic impedance for the system.

(6) TYPICAL MEASUREMENTS ON TRANSISTORS

A few typical results of measurements on various high-frequency transistors are given in Figs. 3 and 4. These show the external characteristics, i.e. no corrections have been made for the internal parasitic elements or measuring resistors. The

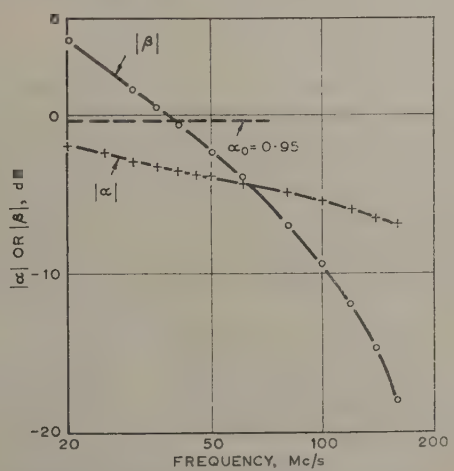


Fig. 3.—Current gain of transistor type V15/20.
$I_e = 4$ mA; $V_c = -6$ volts.

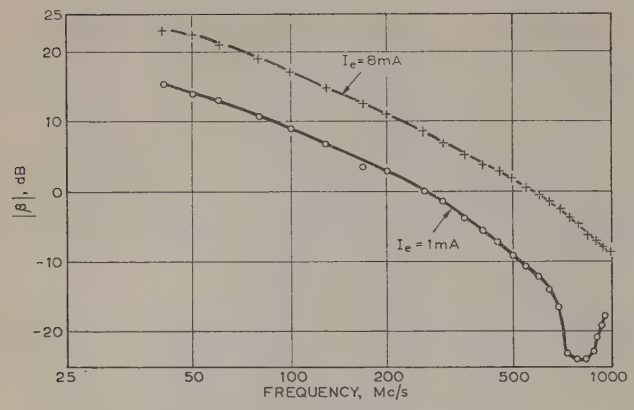


Fig. 4.—Current gain $|\beta|$ of transistor type 2N700.
$V_c = -5$ volts.

effects of, and corrections for, these elements have been discussed at length in the literature.[2, 8] For the reasons given above, most attention has been devoted to $|\beta|$ measurements.

(7) CONCLUSIONS

By the use of transmission-line techniques and attenuators matched to the measuring resistors, the conventional method of f_1 and f_α measurement has been extended to at least 1 Gc/s, with good accuracy for the f_1 measurement.

The two major factors determining the frequency limit of this apparatus are the errors in the measuring arms due to the d.c. decoupling networks and the correction required to take into account the measuring resistances and the header capacitance. Other factors, such as the rise in zero correction at high frequencies, are due to the imperfections of the ancillary equipment available rather than to any fundamental difficulty.

An improvement in the first factor could be made by introducing the bias supplies through the measuring resistors.[3] This would also enable high biasing currents to be used; at present these are limited by the low dissipation of the decoupling resistors.

It is concluded that the techniques described can be extend to cover all frequencies for which transistors mounted in co ventional headers are likely to be used.

(8) ACKNOWLEDGMENTS

The author wishes to thank Mr. L. Bowden and Mr. R. Key of the S.E.R.L. Drawing Office and Workshop, respectively, f their careful work on the measuring jig. Acknowledgment made to the Admiralty for permission to publish the paper.

(9) REFERENCES

(1) DAS, M. B., and BOOTHROYD, A. R.: 'Measurement Equivalent Circuit Parameters of Transistors at V.H.F. *Proceedings I.E.E.*, Paper No. 3077, May, 1959 (**106** Suppl. 15, p. 536).

(2) HYDE, F. J.: 'The Current Gains of Diffusion and Dri Types of Junction Transistors', *ibid.*, Paper No. 293 May, 1959 (**106** B, Suppl. 17, p. 1046).

(3) BASSETT, H. G.: 'An Apparatus for the Measurement Current Gain in the Range 1–300 Mc/s', *ibid.*, Pap No. 2949, January, 1960 (**106** B, Suppl. 15, p. 532).

(4) O'CONNELL, J. H., and SCOTT, T. M.: 'Measurement Transistor Characteristics in the 3–250 Mc/s Frequenc Range', *RCA Review*, 1958, **19**, p. 598.

(5) TERMAN, L.: 'A High-Frequency Alpha Cut-Off Jig', *Tran actions of the Institute of Radio Engineers*, 1959, ED- p. 242.

(6) CRIPPS, L. G.: 'Transistor Cut-Off Frequency Measurement *Proceedings of the Institute of Radio Engineers*, 1958, 4 p. 781.

(7) THOMAS, D. E., and MOLL, J. L.: 'Junction Transistor Shor Circuit Current Gain and Phase Determination', *ibi* p. 1177.

(8) CRIPPS, L. G.: 'Transistor High-Frequency Parameter f *Electronic and Radio Engineer*, 1959, **36**, p. 341.

(9) PAGE, D. F., and BOOTHROYD, A. R.: 'A Simple Measureme for Transistor Current Gain in Magnitude and Phas *Proceedings of the Institute of Radio Engineers*, 1959, 4 p. 1273.

.375.9

The Institution of Electrical Engineers
Paper No. 3224 E
May 1960

COMPARISON OF GAIN, BANDWIDTH AND NOISE FIGURE OF VARIABLE-REACTANCE AMPLIFIERS AND CONVERTORS

By J. D. PEARSON, M.Sc., and J. E. HALLETT, Graduate.

(*The paper was first received 25th June, and in revised form 2nd December*, 1959.)

SUMMARY

Formulae are derived for the gain-bandwidth products and noise ures of a variable-reactance amplifier and a convertor. It is shown t for equal gains and noise figure the convertor has a greater ndwidth than the amplifier. This is confirmed experimentally with ircuit operated as a straight amplifier and a convertor.

LIST OF PRINCIPAL SYMBOLS

f_1, f_2, f_3 = Signal, idler and pump frequencies.
$\omega_1, \omega_2, \omega_3$ = Angular frequencies of signal, idler and pump.
$\omega_1', \omega_2', \omega_3'$ = Resonant angular frequency of signal, idler and pump circuits.
i_i = Current flowing at ω_i.
a_0 = Zero-voltage capacitance of the variable-capacitance diode.
a_1 = Change in capacitance for unit voltage swing.
v_i = Voltage across the diode at ω_i.
L_i = Inductance of the resonant circuit at ω_i.
C_i = Capacitance of the resonant circuit at ω_i.
R_i = Resistance of the resonant circuit at ω_i.
Y_i = Total admittance of the circuit at ω_i.
G_{Ti} = Total conductance at ω_i.
G = Negative conductance.
G_g = Conductance of generator.
G_L = Conductance of load.
G_i = Conductance of the circuit at ω_i.
Q_1, Q_2 = Loaded Q-factors at ω_1' and ω_2'.
Q_{20} = Unloaded Q-factor at ω_2'.
2δ = Bandwidth to 3 dB points.
K = $(\text{Gain})^{1/2} \times$ (fractional bandwidth).
i_{n1} = Noise current at ω_1'.
i_{n2} = Noise current at ω_2'.
k = Boltzmann's constant.
T_0 = Temperature of source.
T_L = Temperature of load.
T = Temperature of circuit.
N = Noise figure.
$\alpha = G_2/G_L$ for a convertor.

The suffix i takes the value 1, 2 or 3 throughout the paper.

(1) INTRODUCTION

The paper discusses the use of a semiconductor diode as the ive element in a parametric amplifier.

Semiconductor diodes possess the property that under conions when the conduction current is small (negative bias and all positive bias) the capacitance of the diode is a function of e applied voltage. The type of amplifier to be considered is e in which a voltage at frequency f_3 is applied to a diode, thus causing a variation in its capacitance at this frequency. Associated with the diode are two resonant circuits at frequencies f_1 and f_2, and f_3 is chosen such that

$$f_1 + f_2 = f_3 \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

It has been shown that the effect of the variation in capacitance at f_3 is to introduce a negative conductance into the f_1 and f_2 circuits. Thus gain is possible at f_1 or f_2, and if a sufficiently large negative conductance is produced, oscillations at f_1 and f_2 occur. We define a system in which a signal enters at f_1, which is also the output frequency as a straight amplifier, and a system in which the signal enters at f_1 and the output is taken at f_2 as a negative-resistance convertor.

The basic theory of this type of amplifier has been given by Heffner and Wade[1] and by Chang and Bloom.[2] These papers did not, however, consider the overall noise figure of a system using a parametric amplifier as its first stage. In conventional valve amplifiers, thermal noise in the load resistor is not amplified by the valve because of its unidirectional properties, and thus this load noise contributes only to the noise in the next stage, which with a high-gain amplifier is insignificant in the overall noise performance. Thus, in the definition of noise figure of a valve amplifier, the noise arising in the load is excluded. However, with a parametric amplifier this exclusion leads to a very optimistic estimate of the noise figure, since, as the parametric amplifier in its simplest form is a bidirectional device, the overall noise figure is increased by the amplified noise arising in the load resistor.

The overall noise contribution of the first stage is considered from the definition of noise figure as the

$$(N_{out}/N_{in}) \div (\text{signal out})/(\text{signal in})$$

the noise out being the total noise in the output load of the parametric amplifier.

It is possible to remove the noise contribution due to the load by the use of circulators, as shown by Siegman,[3] who also pointed out that load noise can be reduced at the expense of bandwidth. Use of a balanced pair of amplifiers with a hybrid-T junction has also been shown[4] theoretically to eliminate load noise.

In Section 2 the theoretical noise figure and gain-bandwidth products of the straight amplifier are compared with the negative-resistance convertor, and it is shown that, for equal noise figures in the two cases, the convertor can have a considerably greater gain-bandwidth product.

In Section 3, experimental results are given for a 40 Mc/s amplifier used both as a straight amplifier and as a negative-resistance convertor. The diode and resonant circuits were the same in both cases, but the form of the external loading was changed to give outputs at 40 Mc/s (straight amplifier) or at 90 Mc/s (convertor). The pump frequency in each case was 130 Mc/s. The results show a much better noise figure and gain-bandwidth product for the convertor.

Written contributions on papers published without being read at meetings are ted for consideration with a view to publication.
Mr. Pearson and Mr. Hallett are with Ferranti, Ltd.

(2) THEORETICAL COMPARISON OF AMPLIFIER AND CONVERTOR

Expressions are derived in Section 7 for the gain-bandwidth product and noise figure for the straight amplifier and the negative-resistance convertor. In this Section they will be put in a form suitable for comparing the performance of the two cases.

We will consider noise figures for high gains and under conditions where $T = T_L = T_0$ for simplicity. Consider first the straight amplifier; we define γ such that

$$G_L = \gamma G_g \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The negative conductance, G, is given by

$$G = \frac{\omega_1'\omega_2' C_p^2}{4G_{T2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

For a straight amplifier the idler circuit is normally unloaded; thus $G_{T2} = G_2$.

It is assumed that C_p is the maximum amplitude of capacitance swing available from the active element. It is also assumed that C_p, G_1 and G_2 are the same in the straight amplifier and convertor, i.e. the same device is being used with different types of external loading.

We define x such that

$$G = xG_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

The quantity x is the ratio of the negative conductance available to the cold losses in the unloaded signal circuit. From eqns. (2), (4) and (54),

$$G_g \simeq \frac{G_1(x-1)}{1+\gamma} \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Substituting from eqns. (2) and (5) into eqn. (55),

$$N_a \simeq \frac{x(1+\gamma)\omega_3}{(x-1)\omega_2'} \quad . \quad . \quad . \quad . \quad . \quad (6)$$

If we are to obtain low noise figures in the straight amplifier it is evident from eqn. (53) that the signal circuit must be considerably loaded. The Q-factor of the signal circuit, for a low-noise amplifier, will therefore be considerably below that of the unloaded idler tank, and thus, for high gains, eqn. (41) reduces to

$$K_a \simeq \frac{2(G_g G_L)^{1/2}\omega_2'}{(G_1 + G_L + G_g)\omega_1' Q_{20}} \quad . \quad . \quad . \quad (7)$$

where Q_{20} is the unloaded Q-factor of the idler tank. Substituting from eqns. (2) and (4) into eqn. (7),

$$K_a \simeq \frac{2\gamma^{1/2}(x-1)\omega_2'}{(1+\gamma)x\omega_1' Q_{20}} \quad . \quad . \quad . \quad . \quad (8)$$

We will now consider the negative-resistance convertor. We define

$$\alpha = G_2/G_L \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Thus $$G_{T2} = G_2 + G_L = G_2(1+\alpha)/\alpha \quad . \quad . \quad . \quad (10)$$

and in the convertor the maximum available negative conductance becomes

$$G = \frac{xG_1\alpha}{1+\alpha} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

The Q-factor of the idler circuit is reduced by the load and becomes

$$Q_2 = \frac{Q_{20}\alpha}{1+\alpha} \quad . \quad . \quad . \quad . \quad . \quad (12)$$

From eqns. (11) and (57),

$$G_g = G_1\frac{(\alpha x - 1 - \alpha)}{1+\alpha} \quad . \quad . \quad .$$

Substituting from eqn. (13) into eqn. (59),

$$N_c \simeq \frac{x\alpha\omega_3}{(\alpha x - 1 - \alpha)\omega_2'} \quad . \quad . \quad .$$

In obtaining the simplified expression (8) for the gain-ban product of the straight amplifier, it was assumed that Q_2 which for low-noise amplifiers will be true. In the co case, however, both the ω_1' and ω_2' circuits are loaded ext Nevertheless, this assumption will be made: it may not b in particular amplifiers if, for instance, noise figure is sa for bandwidth.

Eqn. (42) then becomes

$$K_c \simeq 2\left(\frac{\omega_2'}{\omega_1'}\right)^{3/2}\left(\frac{G_L G_g}{G_{T2}G_{T1}}\right)^{1/2}\frac{1}{Q_2} \quad . \quad .$$

Substituting from eqns. (9), (12) and (13) into eqn. (15),

$$K_c \simeq 2\left(\frac{\omega_2'}{\omega_1'}\right)^{3/2}\frac{1}{Q_{20}}\left[\frac{(1+\alpha)(\alpha x - 1 - \alpha)}{x\alpha^3}\right]^{1/2}$$

Expressions for gain-bandwidth products and noise figur now been obtained for a straight amplifier and a ne resistance convertor having the same swing in capacita the diode and using the same unloaded tank circuits. Th bandwidth products will now be compared when the l has been adjusted in the two cases so that the noise are equal. Thus, if $N_c = N_a$, from eqns. (6) and (14) w

$$\alpha = \frac{1+\gamma}{\gamma(x-1)} \quad . \quad . \quad . \quad . \quad .$$

The gain-bandwidth product of the convertor having the noise figure as the amplifier is obtained by substituting eqn. (17) into eqn. (16):

$$K_c \simeq 2\left(\frac{\omega_2'}{\omega_1'}\right)^{3/2}\frac{(x-1)}{Q_{20}}\left[\frac{\gamma(1+\gamma x)}{(1+\gamma)^3 x}\right]^{1/2} \quad .$$

From eqns. (16) and (18) it follows that

$$\frac{K_c}{K_a} \simeq \left(\frac{\omega_2'}{\omega_1'}\right)^{1/2}\left[\frac{x(1+\gamma x)}{1+\gamma}\right]^{1/2} \quad . \quad .$$

Hence an increased bandwidth results if ω_2'/ω_1' is made Since $x > 1$, the second term is greater than unity, and fo x, obtained by having a negative conductance large com with the unloaded conductance, this term can give rise large increase in bandwidth.

(3) EXPERIMENT

In order to check the theory given above, a circuit was which without external loading exhibited three resonances that $\omega_1' + \omega_2' = \omega_3'$. The resonances ω_1' and ω_2' were c so that they lay within the bands of an available tele f.m. turret tuner which had a noise figure of 7½ dB. The c shown in Fig. 1, consists of a section of coaxial line circuited at each end, with a variable-capacitance diode across it. The coaxial line had an outer conductor of inner diameter and an inner conductor of 0·25 in outer dia In such a system the resonant frequencies are given by

$$\cot\frac{\omega l_1}{c} + \cot\frac{\omega l_2}{c} = \omega C Z_0 \quad . \quad .$$

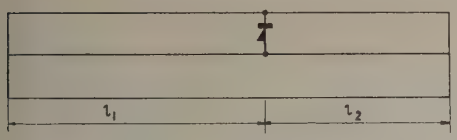


Fig. 1.—Resonant coaxial-line system.

where l_1 and l_2 are the distances of the diode from the short-circuited ends, C is the added capacitance and Z_0 is the characteristic impedance of the line.

The diode used—a 7-volt reference diode—had a capacitance of 24 pF measured on a 1 Mc/s bridge. The encapsulation was changed to eliminate the inductances associated with the fine wires normally used. With $l_1 = 175{\cdot}6$ cm and $l_2 = 125{\cdot}4$ cm, resonances occur theoretically at 42·5 Mc/s, 96 Mc/s and 138·5 Mc/s, and these were observed experimentally with the unloaded resonator.

The theoretical distribution of voltage along the line is shown in Fig. 2.

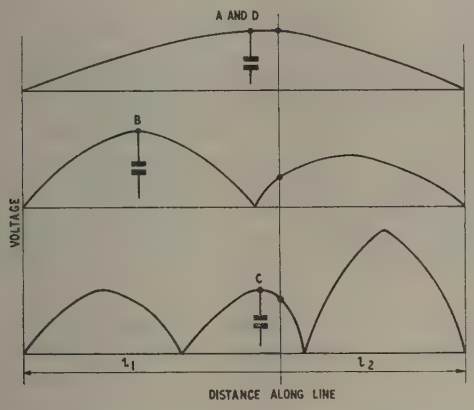


Fig. 2.—Theoretical distribution of voltage at resonance.

The signal input was fed in at A via a capacitor, which theoretically has zero loading of the 96 Mc/s resonance. The output of a negative-resistance convertor was taken at B via a capacitor at a point of maximum coupling to the second resonance.

The pump power in all cases was fed in at C via a 6·8 pF capacitor.

(3.1) Straight Amplifier

During the operation of the amplifier, certain additions were found necessary and were included as difficulties arose. The final arrangement is shown in Fig. 3. The low-pass filter No. 1 had a cut-off frequency of 60 Mc/s and was used to prevent loss of pump power to the untuned signal-source impedance. The low-pass filter No. 2 had a cut-off frequency of 100 Mc/s and also prevented loss of pump power. The pump source had a tuned output circuit and thus did not load the resonator appreciably at frequencies other than the pump frequency. Flexible coaxial cable of 72 ohms impedance was used to join the various components shown.

The signal source was matched to the cable. The detector was frequency dependent, and its admittance together with that of the low-pass filter No. 2 at F is shown in Fig. 4. To obtain the amount of external load, the admittance into the amplifier at E was measured for different coupling capacitances. The

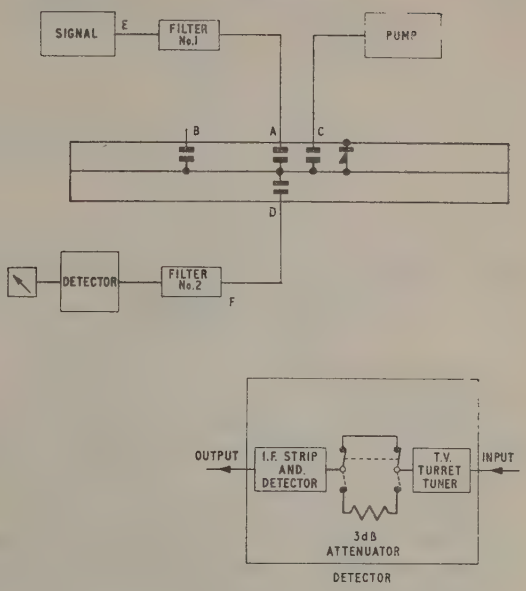


Fig. 3.—Experimental arrangement of straight amplifier.

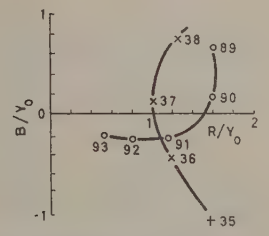


Fig. 4.—Admittance at input of turret tuner.

source impedance was matched to the cable; at resonance the voltage standing-wave ratio is a maximum, and at E

$$(\text{v.s.w.r.})_{Emax} = G_1 + \frac{G_L}{G_g} \quad . \quad . \quad . \quad (21)$$

To measure the gain-fractional-bandwidth product, the turret tuner was adjusted to a given input frequency. The pump supply was then adjusted in frequency to give a known signal gain at a minimum pump power level. The 3 dB bandwidth was then measured. This operation was repeated at various gains and for different input frequencies.

To measure the noise figure of the amplifier, a very small coupling capacitor was inserted at B, Fig. 3, and connected to the signal generator. A matched noise source was connected at E. Initially the noise source was turned off and the pump adjusted as described above to give amplification of the input signal injected at B. The bandwidth was measured and thus, from the gain-bandwidth results already obtained, the gain at this pump setting of a signal injected at E could be estimated. The signal input was then reduced to zero, the pump setting being left fixed. A 3 dB pad was switched into the circuit before the final detector, and the noise generator output was then increased to obtain the original level on the output meter. This was repeated for various gains and input frequencies. This rather complicated setting-up procedure was adopted to avoid errors obtained when a signal source was changed to a noise source

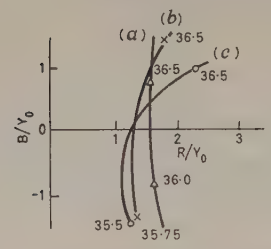


Fig. 5.—Admittance at E (Fig. 3) with various coupling capacitances.
(*a*) Input and output, 18 pF.
(*b*) Input and output, 22 pF.
(*c*) Input and output, 27 pF.

having slightly different admittances, which resulted in a non-optimum pump setting.

Admittance measurements at E for three values of coupling capacitance are given in Fig. 5. Table 1 summarizes the results for these cases. The theoretical noise figures based on measured values of loading are given.

With increased coupling of the input and output circuits (27 pF capacitors) the gain of the amplifier could not be increased above 10 dB for any setting of the pump source. It was assumed therefore that the maximum permissible pump voltage swing had thus been reached.

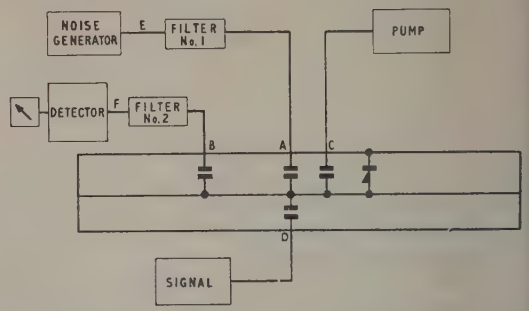


Fig. 6.—Experimental arrangement of negative-resistance converto

frequency sensitive (Fig. 4). In order to estimate the theoreti noise figure, the admittance was measured looking into t amplifier at E for different coupling capacitances. These a plotted in Fig. 7.

We have $$(\text{v.s.w.r.})_E = G_1/G_g$$

and $$N = (1 + G_1/G_g)(1 + \omega_1'/\omega_2')$$

The gain-bandwidth products were measured at differe input frequencies by adjustment of the pump power as with t

Table 1

Fig. 5	C_{in}	C_{out}	f_1 (resonant)	f_1	N		K (measured)		
					(theory)	(measured)	10 dB	15 dB	20 dB
	pF	pF	Mc/s	Mc/s			$\times 10^{-2}$	$\times 10^{-2}$	$\times 10^{-2}$
(*a*)	18·0	18·0	36·25	36·7	3·61	4·3	0·95	1·05	1·2
(*b*)	22·0	22·0	36·15	36·4	3·30	4·5	0·905	0·932	1·06
(*c*)	27·0	27·0	35·9	36·0	3·15		1·2		

Table 2

Fig. 7	C_{in}	C_{out}	f_1 (resonant)	f_1	N		K (measured)		
					(theory)	(measured)	10 dB	15 dB	20 dB
	pF	pF	Mc/s	Mc/s			$\times 10^{-2}$	$\times 10^{-2}$	$\times 10^{-2}$
(*a*)	10·0	2·7	38·3	38·3	2·02	2·24	1·62	1·7	—
(*b*)	15·0	2·7	38·2	38·4	1·82	2·1	1·81	2·15	2·24
(*c*)	18·0	2·7	37·75	37·7	1·69	2·13	—	2·66	2·92
(*d*)	22·0	2·7	37·0	37·5	1·63	2·26	—	3·13	—
(*e*)	27·0	2·7		37·0	1·59	2·4	—	3·01	—

The measured noise figures and gain-bandwidth products given in Table 1 are optimum figures for the given coupling capacitances. The frequency used is slightly greater than the measured resonant frequency based on Fig. 5. This we suggest is due to the pump voltage changing the effective impedance of the variable-capacitance diode.

The noise figure was measured at different values of gain. The quoted noise figures are those obtained at gains sufficiently high for the noise figure to be unaffected by gain variations. Thus the quoted noise figures are the total noise figure of the parametric amplifier alone.

(3.2) Negative-Resistance Convertor

The convertor is shown in Fig. 6. The signal source was matched to the cable, but the admittance of the detector was

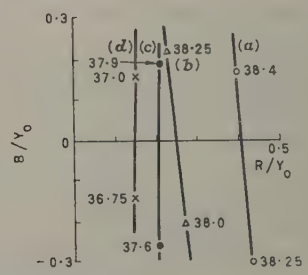


Fig. 7.—Admittance at E (Fig. 6) with various input coupling capacitances and an output coupling capacitance of 2·7 pF.
(*a*) Input, 10 pF. (*b*) Input, 15 pF.
(*c*) Input, 18 pF. (*d*) Input, 22 pF.

aight amplifier. Table 2 summarizes the results obtained for convertor.

With an input capacitor of 27 pF and an output one of 2·7 pF, ns above 16 dB could not be obtained. We assume therefore t the maximum variation of capacitance was not sufficient produce a negative conductance large enough to overcome the ded conductance of the unpumped circuit. The voltage oss the diode was measured on a valve voltmeter under the ximum gain conditions for this case and found to be 0·7 volt ak).

(4) CONCLUSIONS

Experimentally it has been found that a greater gain-band-dth product and lower noise figure may be obtained with a gative-resistance convertor than with a straight amplifier. The oretical noise figure in the case of the convertor showed good reement with the theoretical figure at the lower values of ading; however, with increased loading the noise figure teriorated. It is suggested that the very high swings used for large loadings makes the contribution of other noise sources g. shot noise in the diode) significant in comparison with the hnson noise sources

In the amplifier case the measured noise figure was significantly gher than predicted and deteriorated when the coupling capaci-nces were greater than 18 pF.

The measured results are not conclusive since good agreement th theory was not obtained. Since we might expect that noise ising from current flowing in the diode would be proportional the applied voltage, and the two cases have been investigated to a maximum pump voltage, the noise figure which would be otained with a better diode, i.e. when a smaller voltage swing required, would be better for the convertor than the straight nplifier. This low noise figure is also associated with a greater ndwidth.

(5) ACKNOWLEDGMENTS

Acknowledgment is made to the Admiralty for permission to ublish the paper.

(6) REFERENCES

) HEFFNER, H., and WADE, G.: 'Gain, Bandwidth and Noise Characteristics of the Variable-Parameter Amplifier', *Journal of Applied Physics*, 1958, **9**, p. 1321.

) BLOOM, S., and CHANG, K. K. N.: 'Theory of Parametric Amplifiers', *RCA Review*, 1957, **4**, p. 578.

) SIEGMAN, A. E.: 'Gain, Bandwidth and Noise in Maser Amplifiers', *Proceedings of the Institute of Radio Engineers*', 1957, **12**, p. 1737.

) AUTLER, S. H.: 'Proposals for a Maser Amplifier without Non-Reciprocal Elements', *ibid.*, 1958, **11**, p. 1881.

(7) APPENDICES

(7.1) Circuit Equations

Consider a circuit which has resonant angular frequencies ω_1', ω_2' and ω_3'. In their neighbourhood, the circuit may be represented by the lumped-circuit equivalent of Fig. 8.

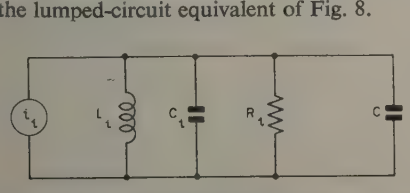


Fig. 8.—Lumped circuit equivalent of parametric amplifier near resonance.

The input signal at $\omega_1 \simeq \omega_1'$ is represented by the current source i_1 and the pump by i_3 (at ω_3). We define

$$\omega_2 = \omega_3 - \omega_1 \quad . \quad . \quad . \quad . \quad . \quad (22)$$

and, since no signal is impressed at ω_2, $i_2 = 0$. We assume that the variation of capacitance with voltage for the diode is such that

$$i = \frac{dq}{dt} = (a_0 + a_1 v)\frac{dv}{dt} \quad . \quad . \quad . \quad . \quad (23)$$

In Fig. 8, the voltage across the circuit may be represented by

$$v = \sum_1^3 v_i' \sin(\omega_i t + \phi_i) = \mathscr{R} \sum_1^3 (-jv_i) \quad . \quad . \quad (24)$$

From eqns. (22), (23) and (24),

$$i_{c1} = ja_0\omega_1(-jv_i) + \tfrac{1}{2}j\omega_1 a_1(-jv_2)^*(-jv_3) \quad . \quad (25)$$

$$i_{c2} = ja_0\omega_2(-jv_2) + \tfrac{1}{2}j\omega_2 a_1(-jv_1)^*(-jv_3) \quad . \quad (26)$$

$$i_{c3} = ja_0\omega_3(-jv_3) + \tfrac{1}{2}j\omega_3 a_1(-jv_1)(-jv_2) \quad . \quad (27)$$

where i_{ci} is the current entering the variable capacitor at ω_i. We now consider the complete circuit at frequency ω_1. We have

$$i_1 = (-jv_1)Y_1 + \tfrac{1}{2}j\omega_1 a_1(-jv_2)^*(-jv_3) \quad . \quad . \quad (28)$$

Similarly

$$0 = (-jv_2)Y_2 + \tfrac{1}{2}j\omega_2 a_1(-jv_1)^*(-jv_3) \quad . \quad . \quad (29)$$

$$i_3 = (-jv_3)Y_3 + \tfrac{1}{2}j\omega_3 a_1(-jv_1)(-jv_2) \quad . \quad . \quad (30)$$

where $$Y_i = \frac{1}{j\omega_i L_i} + j\omega_i C_i + \frac{1}{R_i} + ja_0\omega_i \quad . \quad . \quad . \quad (31)$$

From eqns. (29) and (31),

$$i_1 = (-jv_1)\left(Y_1 - \frac{\omega_1\omega_2 a_1^2|v_3|^2}{4Y_2^*}\right) \quad . \quad . \quad . \quad (32)$$

and from eqns. (29) and (30),

$$i_3 = (-jv_3)\frac{(Y_3 + \omega_3\omega_2 a_1^2|v_1|^2)}{4Y_2} \quad . \quad . \quad . \quad (33)$$

We see from eqn. (33) that the voltage swing at the pump frequency is dependent on signal level, V_1. However, if $|v_3| \gg |v_1|$ we can ignore this dependence and define

$$C_p = a_1 v_3$$

It follows from eqn. (32) that

$$\left(\frac{i_1}{-jv_1}\right) = Y_1 - \frac{\omega_1\omega_2 C_p^2}{4Y_2^*} \quad . \quad . \quad . \quad . \quad (34)$$

We will now consider the case where $\omega_1 = \omega_1'$ and $\omega_2 = \omega_2'$, i.e. the signal and idler circuits are resonant, in which case

$$\left(\frac{i_1}{-jv_1}\right) = G_{T1} - G \quad . \quad . \quad . \quad . \quad (35)$$

where $$G = \frac{\omega_1'\omega_2' C_p^2}{4G_{T2}} \quad . \quad . \quad . \quad . \quad . \quad (36)$$

and G_{T1} and G_{T2} are the total conductance of the circuits at ω_1' and ω_2'. This result was obtained by Heffner and Wade.[1] From eqn. (29),

$$(-jv_2) = -j\omega_2' a_1(-jv_1)^*(-jv_3)/2G_{T2}$$

$$|v_2|^2 = \omega_2'^2 C_p^2 |v_1|^2/4G_{T2}^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (37)$$

(7.2) Gain of a Straight Amplifier

A straight amplifier can be represented by a current source I of source conductance G_g in parallel with a conductance G_1 representing the unpumped losses in the amplifier at the signal frequency, a load conductance G_L and a negative conductance $-G$.

The available power from a current source I of source impedance G_g is $I^2/4G_g$.

The power in the load conductance is

$$I^2G_L/(G_g + G_L + G_1 - G)^2$$

Thus the power gain of the system is

$$\text{gain} = \frac{4G_gG_L}{(G_{T1} - G)^2} \quad . \quad . \quad . \quad . \quad . \quad (38)$$

(7.3) Gain of a Negative-Resistance Convertor

A negative-resistance convertor can be represented at frequency ω_1' by a current source I of source conductance G_g in parallel with a conductance G_1 and a negative conductance $-G$, and at frequency ω_2' by a load conductance G_L in parallel with a conductance G_2, representing the unpumped conductance of the amplifier at ω_2', and a negative conductance $-G'$. The power in the load conductance is $G_L|V_2|^2$. Substituting from eqn. (37), the output power is

$$G_L\frac{\omega_2'^2C_p^2|V_1|^2}{4G_{T2}^2} = G_L\frac{\omega_2'G}{\omega_1'G_{T2}}|V_1|^2$$

$$= \frac{\omega_2'G_LG|i_1|^2}{\omega_1'G_{T2}(G_{T1} - G)^2}$$

from eqn. (35).

Thus,
$$\text{gain} = \frac{\omega_2'}{\omega_1'}\frac{4G_gG_LG}{G_{T2}(G_{T1} - G)^2} \quad . \quad . \quad . \quad . \quad (39)$$

(7.4) Gain-Bandwidth Products

The bandwidth of the straight amplifier and convertor has been shown[1] [eqn. (24)] to be

$$2\delta = \frac{G_{T1} - G}{Q_1G_{T1} + Q_2G\omega_1'/\omega_2'} \quad . \quad . \quad . \quad (40)$$

Thus, from eqns. (38), (39) and (40) for the straight amplifier,

$$K_a = \frac{2(G_gG_L)^{1/2}}{Q_1G_{T1} + Q_2G\omega_1'/\omega_2'} \quad . \quad . \quad . \quad . \quad (41)$$

and for the convertor,

$$K_c = \left(\frac{\omega_2'G_LG_gG}{\omega_1'G_{T2}}\right)^{1/2}\frac{2}{Q_1G_{T1} + Q_2G\omega_1'/\omega_2'} \quad . \quad (42)$$

(7.5) Noise Figure

In deriving an expression for the noise figure, we will consider only Johnson noise in the circuit. Other sources of noise have been considered by Heffner and Wade.[1] They conclude that contributions from sources other than Johnson noise will be small in diode amplifiers. Experimental results given below confirm this, provided that the swing in voltage on the diode is not sufficiently high to cause conduction.

We now calculate the noise output of an amplifier. Consider noise currents i_{n1} and i_{n2} at frequencies ω_1' and ω_2' impressed on the circuit of Fig. 8. Eqns. (28) and (29) are replaced by

$$i_{n1} = (-jv_1)G_{T1} + \tfrac{1}{2}j\omega_1'a_1(-jv_2)^*(-jv_3) \quad . \quad . \quad (43)$$

$$i_{n2} = (-jv_2)G_{T2} + \tfrac{1}{2}j\omega_2'a_1(-jv_1)^*(-jv_3) \quad . \quad . \quad (44)$$

The noise currents are given by

$$|i_{n1}|^2 = 4k\Delta f(G_gT_0 + G_LT_L + G_1T) \quad . \quad . \quad . \quad (4$$

$$|i_{n2}|^2 = 4k\Delta f(G_2T). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4$$

in the straight amplifier, and

$$|i_{n1}|^2 = 4k\Delta f(G_gT_0 + G_1T) \quad . \quad . \quad . \quad . \quad . \quad (4$$

$$|i_{n2}|^2 = 4k\Delta f(G_2T + G_LT_L) \quad . \quad . \quad . \quad . \quad . \quad (4$$

in the negative-resistance convertor.

Eliminating v_2 from eqns. (43) and (44),

$$i_{n1} - j\frac{\omega_1'a_1}{2G_{T2}}(-jv_3)i_{n2}^* = (-jv_1)\left(G_{T1} - \frac{\omega_1'\omega_2'a_1^2|v_3|^2}{4G_{T2}}\right) \quad (4$$

Since the noise currents are uncorrelated $i_{n1}i_{n2}^* = 0$.

Thus, squaring eqn. (49),

$$|i_{n1}|^2 + \frac{\omega_1'^2a_1^2|v_3|^2|i_{n2}|^2}{4G_{T2}^2} = |v_1|^2\left(G_{T1} - \frac{\omega_1'\omega_2'a_1^2|v_3|^2}{4G_{T2}}\right)^2 \quad (5$$

Similarly

$$|i_{n2}|^2 + \frac{\omega_2'^2a_1^2|v_3|^2|i_{n1}|^2}{4G_{T1}^2} = |v_2|^2\left(G_{T2} - \frac{\omega_1'\omega_2'a_1^2|v_3|^2}{4G_{T1}}\right)^2 \quad (5$$

The definition of noise figure is

$$N = \frac{N_{out}}{N_{in}} \div \frac{\text{signal out}}{\text{signal in}}$$

The input noise is defined as the noise output of a resistor at temperature T_0, and it is usual to take $T_0 = 290°\text{K}$.

Thus

$$N = \frac{N_{out}}{kT_0\Delta f \times \text{gain}} \quad . \quad . \quad . \quad . \quad . \quad (52)$$

The output noise for the straight amplifier is equal to $|v_1|^2G_L$. Thus, from eqns. (38), (45), (46) and (52),

$$N_a = 1 + \frac{G_1}{G_g}\frac{T}{T_0} + \frac{G_LT_L}{G_gT_0} + \frac{G\omega_1'T}{G_g\omega_2'T_0} \quad . \quad . \quad (53$$

For high gains, from eqn. (38),

$$G \simeq G_1 + G_g + G_L \quad . \quad . \quad . \quad . \quad . \quad (54$$

and if $T = T_0 = T_L$ eqn. (53) is reduced to

$$N_a = \frac{G_{T1}}{G_g}\left(1 + \frac{\omega_1'}{\omega_2'}\right) = \frac{G_{T1}}{G_g}\frac{\omega_3}{\omega_2'} \quad . \quad . \quad (55)$$

The output noise for the negative-resistance convertor is $|v_2|^2G_L$, and thus, from eqns. (39), (47), (48) and (52).

$$N_c = 1 + \frac{G_1T}{G_gT_0} + \frac{\omega_1'}{\omega_2'}\frac{G_{T1}^2}{GG_{T2}G_g}\left(G_L\frac{T_L}{T_0} + \frac{G_2T}{T_0}\right) \quad (56)$$

For high gains, from eqn. (39),

$$G \simeq G_g + G_1 = G_{T1} \quad . \quad . \quad . \quad . \quad . \quad (57)$$

$$N_c = 1 + \frac{G_1}{G_g}\frac{T}{T_0} + \frac{\omega_1'}{\omega_2'}\frac{G_{T1}}{G_{T2}}\frac{1}{G_g}\left(\frac{G_L}{T_0}T_L + \frac{G_2T}{T_0}\right) \quad (58)$$

and if
$$T = T_0 = T_L$$

$$N_c = \left(1 + \frac{G_1}{G_g}\right)\left(1 + \frac{\omega_1'}{\omega_2'}\right) \quad . \quad . \quad . \quad . \quad (59)$$

$$= \frac{G_{T1}\omega_3}{G_g\omega_2'}$$

621.391.821

The Institution of Electrical Engineers
Paper No. 3158 E
May 1960

A STUDY OF ATMOSPHERIC RADIO NOISE RECEIVED IN A NARROW BANDWIDTH AT 11 Mc/s

By C. CLARKE, Associate Member.

(*The paper was first received* 23rd *May, and in revised form* 8th *September*, 1959.)

SUMMARY

Equipment is described for studying atmospheric radio noise by means of continuous-film photography and by measuring the amplitude probability distributions by a pulse-counting technique. Results are given for observations in England during the summer and autumn of 1956. The observations were made in a bandwidth of 400 c/s.

No simple mathematical representation of the amplitude probability distribution has been found which will fit all conditions, and a graphical presentation is suggested as a means of providing communication engineers with the required data.

The high-frequency radiation consists of a burst of quasi-continuous noise throughout the duration of the flash. This suggests that a burst-duration distribution may be a useful additional measurement in defining the interference effect of the noise.

(1) INTRODUCTION

In radiocommunication the engineer needs a knowledge of atmospheric noise levels in order to assess the radiated power required to maintain a given grade of service. The International Radio Consultative Committee (C.C.I.R.) has published predictions of the noise power to be expected from atmospherics in any part of the world,[1] and these show that atmospherics are often the main source of interference in some frequency ranges, particularly in tropical regions, apart possibly from other stations.

To assess required signal levels it is necessary to assume the signal/noise power ratios, and these are based largely on past experience. For a full understanding of the noise problem, to assess the required signal/noise ratios in future systems and to study means for minimizing the influence of noise, a detailed knowledge of the noise structure is required.

In recent years considerable progress has been made in this field at very low frequencies, 10–100 kc/s, both on a statistical basis[2] and from a study of the wideband waveforms of the field variations from individual lightning discharges.[3, 4] At the low-frequency end of this band, for example, it has been shown that the noise consists largely of discrete impulses originating in the return strokes of lightning discharges. Studies at the Radio Research Station,[5, 6] Slough, show that a reasonably accurate description of this noise may be given in terms of the average level of the noise envelope and the amplitude distribution of the voltage peaks, after passing through a communication-type receiver. A corresponding method of describing noise at higher frequencies is also required.

The variation of the noise level on frequencies between 2·5 and 20 Mc/s has been measured by a subjective technique for many years[7] but the noise structure is known in only very general terms. Japanese workers[8] have measured a few amplitude probability distributions at 3 Mc/s by a technique similar to that described in this paper, and it is known that the noise is less impulsive than at very low frequencies and more nearly approaches the character of thermal noise. Consequently, parameters describing the noise at very low frequencies may not be appropriate at higher frequencies. Preliminary observations were therefore made at the Radio Research Station during the summer and autumn of 1956 on a frequency of 11 Mc/s to assess the merits of the amplitude probability distribution (a.p.d.) as a suitable type of measurement of the noise structure and to determine whether other characteristics need to be measured also.

The investigation was divided into two sections. First the h.f. noise was recorded photographically by means of continuous-film cameras, with provision for simultaneously recording on the same film the noise at very low frequencies from a 10 kc/s cathode-ray direction-finder (c.r.d.f.) display. This permitted a direct comparison of noise at high and very low frequencies, and, on occasions, gave opportunities of relating the h.f. noise to specific storms of known geographical position. Secondly, the a.p.d. of the noise envelope was measured by means of a pulse-counting technique.

The findings of the investigation were embodied in the design of equipment to be used for an extensive programme during the International Geophysical Year, but the preliminary measurements provided useful data in themselves.

(2) EQUIPMENT

A major problem in investigating noise at high frequencies is to find a clear channel, free from interference from station transmissions and yet of a sufficiently wide bandwidth that the measured characteristics of the noise are representative of its effect upon a typical communication system. A compromise is necessary, and for the work described, a communications receiver was used, with a bandwidth of 400 c/s between 3 dB points obtained by a crystal-filter circuit which gave a high attenuation rate to the skirts of the response curve.

Although directional aerials would be used in a practical communication system and would undoubtedly influence the character of the noise, it was thought best for this investigation to use an omnidirectional system to reduce the number of variables under consideration. The aerial consisted of a quarter-wave vertical rod with a counterpoise of eight radial wires, $\lambda/2$ in length, arranged symmetrically around the base of the aerial at ground level.

A schematic representation of the equipment is shown in Fig. 1. The receiver output was taken, at an intermediate frequency of 455 kc/s, to two video-frequency amplifiers, one of which was used for the cathode-ray-tube display and the other for the channel measuring the a.p.d. of the noise envelope. The output from a cathode-ray direction-finder operating at a frequency of 10 kc/s[9] was displayed on a second cathode-ray tube and the traces from both tubes were photographed with one camera. Film-transport speeds normally used were either 10 or 100 in/s.

To find the a.p.d. the fraction of the time that the envelope voltage exceeded a series of thresholds was measured, and this is referred to as the occupation time. This could then be plotted as a function of the threshold to give the distribution.

As shown in Fig. 1, the output from the video-frequency amplifier was detected and the resultant modulation voltage

Written contributions on papers published without being read at meetings are invited for consideration with a view to publication.
The paper is an official communication from the Radio Research Station, Department of Scientific and Industrial Research.

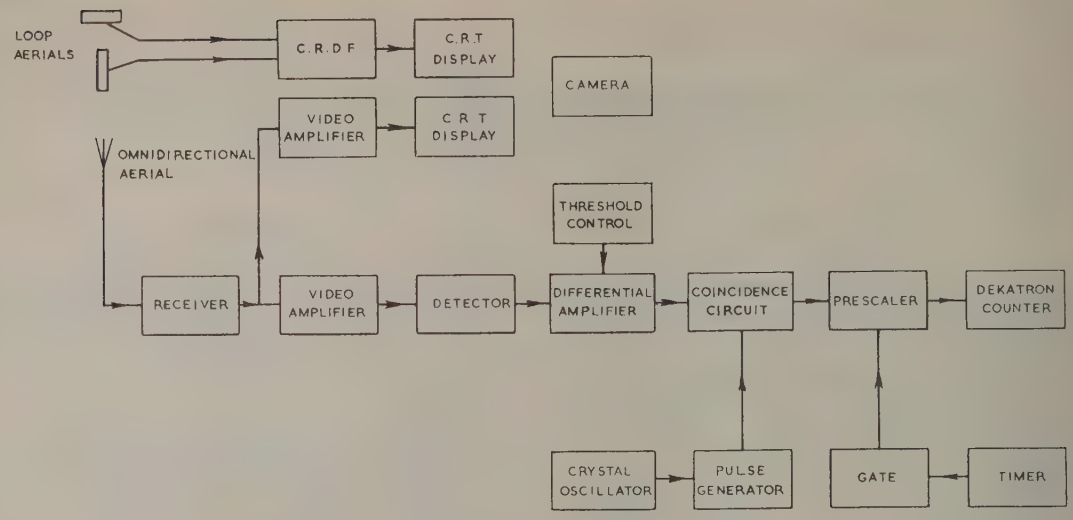


Fig. 1.—Block schematic of the equipment.

passed to a differential amplifier where the threshold was applied. Signals exceeding the threshold were directly coupled to a coincidence circuit where they were used to gate 100 kc/s pulses obtained from a crystal oscillator and pulse generator. Whenever the noise envelope was above the threshold, the coincidence circuit allowed 100 kc/s pulses to pass to a pre-scaler. A gate, controlled from a timing unit, was introduced at this stage and allowed the pre-scaler to pass pulses for a fixed period, usually 10 sec, and the number of 100 kc/s pulses passed in a given ten seconds was recorded by the pre-scaler and a series of Dekatron counters. The ratio of the number generated could then be expressed as a percentage occupation time. In practice, by using 100 kc/s pulses timed for 10 sec, the counters could b made to read percentage directly.

Some details of the analysis circuits are shown in Fig. 2. V is the differential amplifier in which the signal is applied to on grid and the threshold voltage to the other. An output is thu obtained only when the detected signal voltage exceeds th threshold voltage. This signal is directly coupled to the coinci dence circuit composed of V4 and V5. The cathode circuit o V4 is completed by V5A, so that V4 cannot conduct unles V5A is conducting. V5A is normally non-conducting, but i switched on by the 100 kc/s pulses applied to the grid and thes pulses then appear at the cathode of V4. In the no-signal con

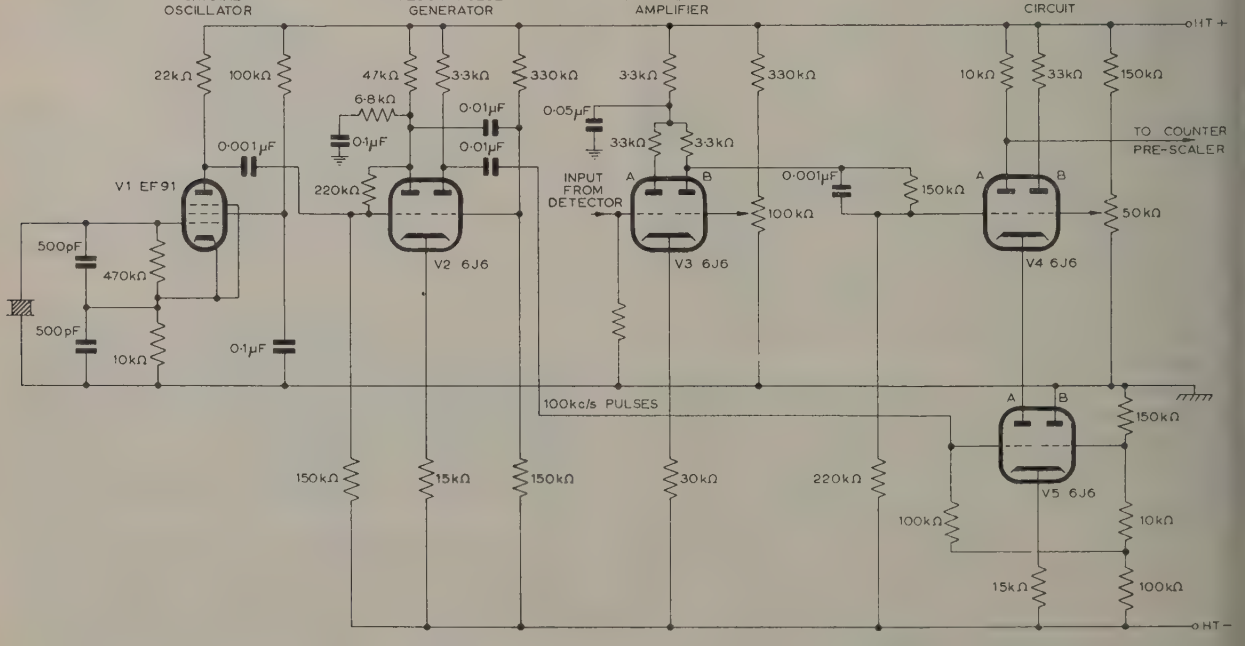


Fig. 2.—Detailed circuit of threshold and clock-pulse units.

ion the grid potential of V4B is adjusted so that these pulses t fail to operate the pre-scaler connected to the anode of V4A. henever the noise voltage exceeds the threshold, V4A is itched on and 100 kc/s pulses appear at the anode and operate e pre-scaler.

The pre-scaler consists of two decade stages of type E1T unters which reduce the pulse frequency to 1 kc/s, suitable the subsequent Dekatron stages. A bistable multivibrator cuit triggered by pulses from an additional bank of Dekatron unters which divide the mains supply frequency and produce output pulse every ten seconds, controls the pre-scaler by ans of a simple diode gate.

(3) OBSERVATION TECHNIQUE

The percentage occupation time was observed, usually between and 1 or 0·1%, as a function of threshold voltage. At low esholds the median of three 10 sec samples was used, but as e threshold was increased the number of samples was also reased to a maximum of 7 or 9 according to the scatter countered among samples. By using the median of a mber of 10 sec samples rather than a single longer period, more representative final reading could be obtained in e presence of intermittent station interference, by monitoring ually and aurally and, if necessary, rejecting contaminated mples.

To relate the threshold voltage to field strength, a calibration is made for each individual run by determining the equivalent ak c.w. voltage from a generator for a given reference eshold, which was then related to the field strength from a owledge of the pick-up factor of the aerial. This was deter-ined experimentally by using a portable transmitter.

For the camera records, the transport speed of 10 in/s per-tted a study of the general character of the noise, and 100 in/s ovided more detailed data on individual noise bursts. The .d.f. information was generally presented as a directional trace assist in relating individual noise bursts to specific storms ose geographical location was known from the British eteorological Office network of direction-finders.[10] Alter-tively, the output from the two channels of the c.r.d.f. could combined in quadrature to give a 10 kc/s noise trace, in a mparable bandwidth to that used at high frequencies, allowing direct comparison of h.f. and v.l.f. noise. Calibrating signals re recorded on each film to relate the noise envelope to an uivalent peak c.w. field as described for the probability stribution.

(4) RESULTS

Cumulative amplitude probability distributions were taken at rious hours during June–November, 1956, and provided liable estimates of the diurnal variation of both the noise vels and the distributions throughout this period.

(4.1) Amplitude Probability Distributions

A typical amplitude probability distribution for quiet summer nditions is shown in Fig. 3. The percentage occupation time, 0 $Q(v)$, is plotted as a function of the threshold voltage, v, hile the ordinate (y) scale is such that the a.p.d. of thermal noise, hich follows the Rayleigh law

$$Q(v) = \exp(-v^2)$$

ould be represented by a straight line at a slope of −2.

The scatter among the 10 sec samples used in this observation riod at any given threshold is illustrated in Table 1, which ows a summary of the results from which this distribution was otted.

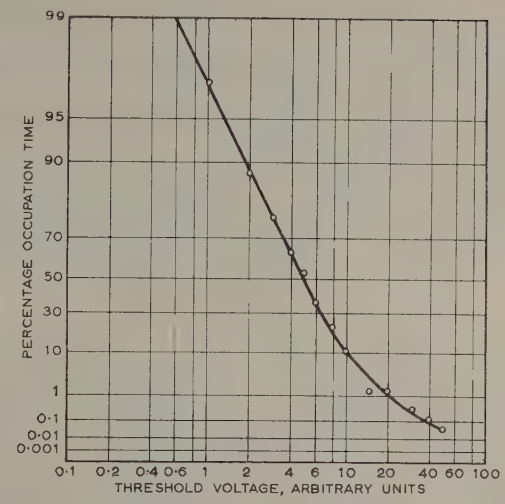


Fig. 3.—Amplitude probability distribution, 0830 h, 20th July, 1956.

Table 1

TYPICAL A.P.D. DATA AT 11 Mc/s

Threshold voltage*	Occupation time, 10 sec sample			Number of samples
	Highest	Lowest	Median	
volts	%	%	%	
0	100			
1	98·3	96·0	97·1	3
2	90·0	85·6	88·4	3
3	82·6	75·3	77·5	4
4	69·2	61·0	64·3	4
5	56·4	44·8	52·9	4
6	44·0	33·4	35·7	4
8	29·3	16·1	21·8	5
10	14·1	5·6	10·1	5
15	2·9	0·56	1·4	5
20	2·0	0·52	1·3	7
30	0·76	0·06	0·30	7
40	0·28	0·02	0·11	7
50	0·07	0·01	0·04	7

* A threshold voltage of 10 volts corresponds to a peak c.w. field of 0·08 μV/m.

Certain parameters that have been directly measured to describe noise in the past may be obtained from the amplitude probability distribution. These include, for example, the average and r.m.s. levels of the noise envelope, which are given by

$$v_{av} = \int_0^\infty Q(v)dv$$

$$v_{rms} = \left[2\int_0^\infty vQ(v)dv\right]^{1/2}$$

In finding the r.m.s. value of the envelope the highest voltage ranges are important, even though they are exceeded for only a small percentage of the time. In certain instances it was necessary to extrapolate the distributions based on experience from analysis of photographic records.

To summarize the results, individual runs were normalized to the 50% occupation-time threshold and the median value of the occupation time at each threshold plotted in 4 h time blocks

from A, representing 0000 to 0400 G.M.T., to F which represents 2000 to 2400 G.M.T.

Parameters deduced from these normalized plots are shown in Table 2. They are as follows:

(*a*) The percentage occupation time (O.T.) above which the a.p.d. may be represented by a Rayleigh law.
(*b*) The amplitude range of the envelope in decibels, represented by the occupation time from 90 to 1%.
(*c*) The ratio in decibels of the r.m.s. to the average voltage of the noise envelope.

This last parameter is perhaps the best single index of noise structure, based on common types of measurement.

Table 2

PARAMETERS DEDUCED FROM COMPOSITE DISTRIBUTIONS

Parameter	Summer time blocks					
	A	B	C	D	E	F
Minimum O.T. for Rayleigh law	60%	50%	20%	35%	70%	60%
Range for 90–1% O.T., dB	25	22·5	22	25	27	26
Ratio r.m.s. to average, dB	3·0	2·5	2·5	3·0	3·5	3·5
	Autumn time blocks					
	A	B	C	D	E	F
Minimum O.T. for Rayleigh law	8%	25%	5%	40%	20%	10%
Range for 90–1% O.T., dB	18·5	19·5	18·5	21	(19)	18
Ratio r.m.s. to average, dB	*	*	1·5	*	*	*

* Insufficient data in high-amplitude regions.

For both summer and autumn conditions the Rayleigh distribution is most closely followed for time block C (08–12 hours G.M.T.) when local thunderstorm activity is low and ionospheric absorption is increasing; the greatest departure occurs in the period before sunset when thunderstorm activity within, say, 2000 km is at a maximum. Thus in summer the distribution curve departs from a Rayleigh law at 70% occupation time for time block E and in the autumn at 40% for time block D.

The mean distributions for the most impulsive summer block E and the smoothest autumn block C are illustrated in Figs. 4(*a*) and (*b*), respectively. It is interesting to compare the dynamic range of the noise represented by these two curves with true Rayleigh noise and also with v.l.f. noise on 10 kc/s as quoted by Horner and Harwood.[5] The figures are shown in Table 3 normalized to the r.m.s. level, and they emphasize the similarity between thermal noise and h.f. atmospheric noise for the autumn block C and the large departure from thermal noise of v.l.f. atmospherics.

Table 3

COMPARISON OF ATMOSPHERIC AND THERMAL NOISE DISTRIBUTIONS

Occupation time	Threshold relative to r.m.s. level			
	10 kc/s	11 Mc/s Summer, E	11 Mc/s Autumn, C	Thermal noise
	dB	dB	dB	dB
90%	·32	-17	−10·5	−9·9
50%	20	--7·5	−2	−1·5
Average	12	--3·5	−1·5	--1·1
R.M.S.	0	0	0	0
5%	−1	6·5	5	4·7
1%	11	11	8	6·8
0·1%	24	17	14	8·4

(4.2) Diurnal Variation of Noise Levels

The diurnal variation of the noise level may be expressed in terms of the median level of the envelope, i.e. the threshold voltage exceeded by the noise envelope for 50% of the time. In addition, an appreciation of the interference effect of the noise may be obtained by combining the median results with other percentile values, say 10 and 1%.

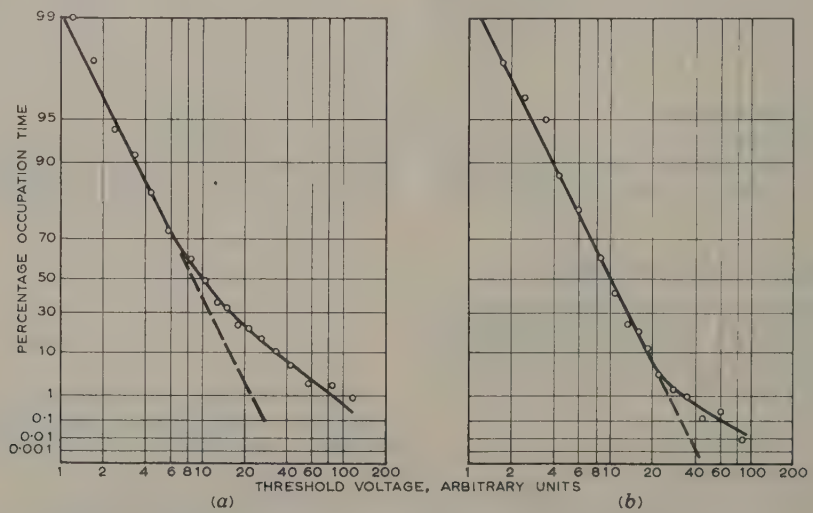


Fig. 4.—Composite amplitude probability distributions of the noise envelope.
(*a*) Summer, 1600–2000 h.
(*b*) Autumn, 0800–1200 h.

he equivalent peak c.w. field exceeded by the noise envelope 50, 10 and 1% of the time, in the bandwidth of 400 c/s, is wn in Figs. 5(*a*) and (*b*) for the summer and autumn con- ons, respectively. In summer the median level shows a nal variation of about 12 dB and the 10 and 1% levels show B. The ratio of the 10 and 1% values to the median level eases in the afternoon and evening because the noise becomes e impulsive in character. In autumn the diurnal variation bout 10 dB for all three percentile values, the quieter autumn ditions with little local activity producing a less impulsive e structure than in summer.

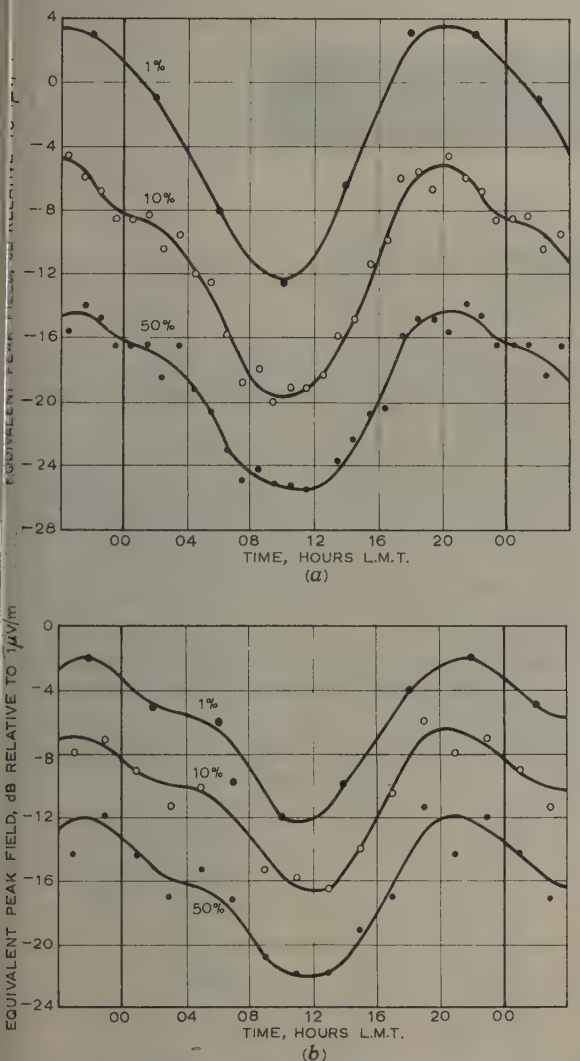


Fig. 5.—Diurnal variations of the amplitude of the noise envelope exceeded for 50, 10 and 1% of the time.

(*a*) Summer, 1956.
(*b*) Autumn, 1956.

(4.3) Field from a Single Atmospheric Source

One use of the film records was to examine the field radiated om discrete distant sources of atmospherics, whose geo- aphical locations were known, in an attempt to explain certain atures of the distribution curves. The following results were obtained when an isolated storm was in progress off Southern Sardinia, at a range of 1 600 km, between 14 and 15 h L.M.T. in December, 1955. A number of well-defined atmospheric noise bursts were obtained, correlating with the appropriate c.r.d.f. bearing, and their amplitude probability distributions were measured from the film.

Each individual burst approximated very closely to a thermal noise distribution and had a mean value of 1μV/m for the field exceeded by the noise envelope for 50% of the time. Since the burst amplitudes were Rayleigh distributed, the r.m.s. value of the envelope is 1·5 dB above this figure.

(5) DISCUSSION OF RESULTS

(5.1) Source Effects

The film records have also been used to study the relationship between the noise envelope at high frequencies and the corresponding signals at 10 kc/s received on the cathode-ray direction-finder. Horner[11] has already described some preliminary results on this aspect, and some additional work has been described by Horner and the author.[12]

It has been suggested that the main source of h.f. atmospheric noise arises from the precursor discharge or stepped-leader stroke. Aiya[13] has measured the interference effect of atmospherics to broadcasting services in India, and suggests that the noise levels may be explained almost entirely on the assumptions that:

(*a*) The majority of atmospheric noise in the h.f. band originates in cloud discharges.

(*b*) Each discharge contains four stepped-leader strokes.

The film records taken at Slough during 1956 show that, as is already established, radiation at very low frequencies consists of discrete impulses originating in individual ground strokes. On 11 Mc/s, however, the radiation associated with both ground and cloud discharges may consist of an almost continuous burst of noise for each flash.

This is illustrated in Figs. 6(*a*), (*b*) and (*c*) in which are displayed, on the top trace of each record the output from the cathode-ray direction-finder on 10 kc/s, and on the bottom trace, noise at 11 Mc/s. In addition, Figs. 6(*a*) and (*b*) show a third trace between the other two, which is the output from a wide-band (5–25 kc/s) aperiodic receiver (waveform recorder). These waveform traces are ignored in the present discussion. The 11 Mc/s trace represents the undetected intermediate-frequency output from the receiver, in a 400 c/s bandwidth; the lower half of the symmetrical trace has been masked to obtain a simpler record. The c.r.d.f. trace shows directional records obtained at a frequency of 10 kc/s and with a bandwidth of 250 c/s. No masking has been used on this trace.

Fig. 6(*a*) shows the quasi-continuous nature of the 11 Mc/s signal, and the very complex nature of the c.r.d.f. traces suggests that it was a close flash—possibly a cloud discharge. The overall duration is nearly 500 millisec. Preceding this complex discharge is a short burst of noise of about 75 millisec duration, the termination of which coincides in time with a c.r.d.f. trace. This type of correlation is often observed and is interpreted in this case as a distant flash, the noise coming from the discharge mechanism prior to the return stroke which gives rise to the very-low-frequency c.r.d.f. trace.

Fig. 6(*b*) taken from the same film shows a similar quasi-continuous type of discharge at 11 Mc/s, but the c.r.d.f. record shows a series of discrete traces, all on the same bearing, representing individual strokes of a discharge to earth. The overall duration is between 600 and 700 millisec, including a burst of noise of about 30 millisec duration before the first return stroke.

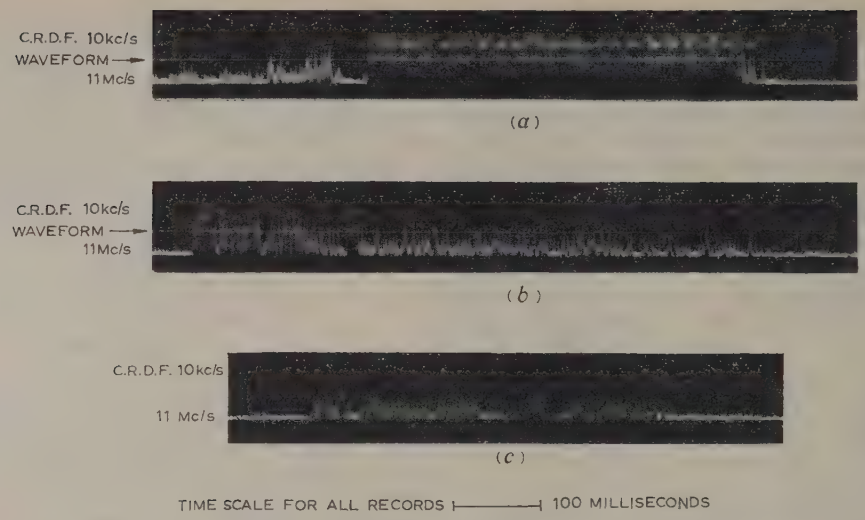


Fig. 6.—Radiation at 11 Mc/s and 10 kc/s from individual flashes.
(*a*) From close cloud discharge.
(*b*) From close ground discharge.
(*c*) From distant ground discharge.

Fig. 6(*c*) shows a discharge of about 450 millisec duration which was, in fact, radiated by a single flash from the Mediterranean storm discussed in Section 4.3. From the peak amplitude of the c.r.d.f. record it is reasonable to assume that this was a ground discharge.

It is obvious that in these examples the stepped-leader mechanism contributed little to the energy radiated at high frequencies because it is generally agreed that the duration for the actual stepped part of the discharge is comparatively short, say 10 millisec. Also, for discharges to earth and possibly for cloud flashes, more than one stepped process is unusual, as it implies deionization of the channel. Continuous-type noise up to 100 millisec duration has been observed before the first return stroke, as in Fig. 6(*b*), and this probably arises from the precursor discharge observed by Clarence and Malan.[14] The more typical appearance of high-frequency radio noise, however, is that of a substantially continuous process of radiation, present throughout the duration of the flash, the precise mechanism for which is not yet fully understood, but which probably includes radiation arising from the J-streamer components described by Malan and Schonland.[15]

(5.2) Propagation Effects

One difficulty when considering the diurnal variation of the noise field strength is to decide whether the variation is caused mainly by propagation conditions or whether it is more directly related to the incidence of thunderstorms. To illustrate this effect, the diurnal variation of the field corresponding to the median level of the noise envelope during the summer of 1956 is plotted in Fig. 7, together with the E-layer critical frequency and the F-layer skip distance, taking mean values for the season.

The diurnal variation in the noise level shows some inverse correlation with f_0E, which suggests that the main factor influencing the noise level may be ionospheric absorption. Examining the curve in more detail, between 0400 and 0800 h there is a steady reduction in skip distance from about 2 000 to 1 000 km but the noise level is also steadily decreasing. Thus, although the possible effective thunderstorm area is increased, at that time of day there is little activity in the additional include area and the absorption is the main factor controlling th noise level.

Between 1200 and 1800 h the noise level rises from i minimum to near its maximum value. The rise is more rapi than would be expected from a simple reduction in ionospheri absorption and as the skip distance remains constant, it suggest that for the latter part of this period the controlling factor ma be thunderstorm activity at ranges in excess of 1 000 km. Clos activity may also be a contributory factor, with propagation vi sporadic-E ionization.

Beyond, say, 2200 h, there is a gradual reduction in nois level as the thunderstorm activity decreases and the skip distanc increases until ionospheric absorption causes the more rapid dro from 0400 h onward.

Summarizing these results it may be said that the main facto affecting the diurnal change in noise is ionospheric absorption

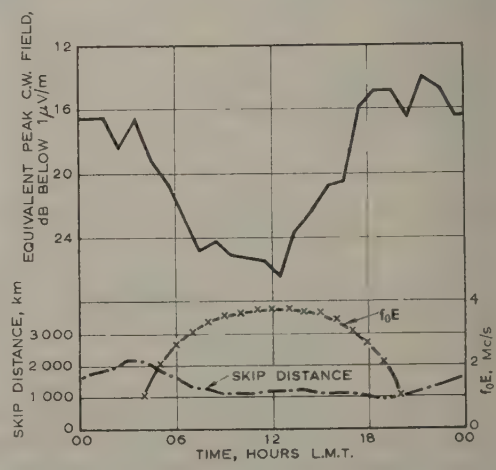


Fig. 7.—Relationship of median-noise field strength to skip distance and f_0E.

t that increased thunderstorm activity in the late afternoon
evening causes a rapid increase in the level. On a seasonal
rage there is no obvious correlation between noise level and
distance but on individual occasions, particularly at night,
ore definite relationship would probably exist.

(5.3) Comparison with Galactic Noise

is interesting to compare the noise levels with those which
have been expected from galactic sources. The lowest
ospheric noise levels in the summer of 1956 occurred in time
ck C, but they are still about 6 dB above the estimated value
galactic noise.[1] In the autumn, although the lowest levels
urred in the same time block, the F2 critical frequency was
erally higher than 11 Mc/s and no galactic noise was received.
nsidering the period 2100 to 0600 h when both absorption and
F2 critical frequency were low, the lowest individual noise
els recorded are comparable with the galactic level but the
sonal median is above this level by 4–8 dB, a figure similar
hat obtained during the summer.

(5.4) Distribution Laws

Attempts have been made in the past to fit empirical formulae
amplitude distributions. Such formulae would have two
vantages. Firstly, if a universal formula could be found, one
two measurements would be sufficient to define the whole
tribution. Secondly, if the formula lent itself to simple
thematical treatment, the derivation of other parameters from
amplitude probability distribution would be simplified.

At very low frequencies a log-normal law has been shown to
ply to a substantial range of the amplitude probability distribu-
n,[16] and such a law applied to the h.f. results for the 'noisy'
mmer time block is shown by the full curve of Fig. 8. In this
gure the ordinate represents occupation time and the abscissa
he log of the threshold voltage V. A log-normal distribution
represented by a straight line which may be defined by the
an logarithm and the standard deviation of log V. It will be
n that a straight line with a slope of $-1{\cdot}4$ gives a fair repre-
ntation of the distribution from 99 to 1% occupation time.
owever, the distribution in this time block departs from
yleigh at 70% occupation time. For other time blocks,
en the distribution is more similar to the Rayleigh type,
log-normal plots show a marked curvature at low thresholds,
g. the dashed curve, which is for autumn block C. The log-
rmal distribution gives a fair representation of the plots from
point of departure from Rayleigh down to 1% occupation
ne, where curvature may again occur, but the data are rather
arse for a positive conclusion.

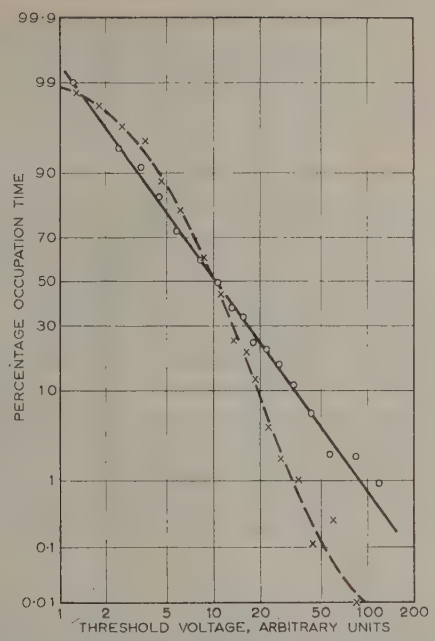


Fig. 8.—Amplitude probability distribution of noise envelope (log-normal presentation).
—— Summer 1600–2000 h.
----- Autumn 0800–1200 h.

From the theoretical point of view there is justification for attempting to apply a Rayleigh law coupled with other parameters to indicate the degree of departure from true fluctuation noise, since atmospheric noise will always approach the Rayleigh distribution at high occupation times. In this region atmospheric noise represents the integrated effect of large numbers of disturbances each of which may contain a number of impulses with random phases and amplitudes. The effect is illustrated in Fig. 9, which shows distributions obtained from a film record and which enables individual sections of the noise to be studied in detail. The Figure shows the high threshold regions of the amplitude probability distributions for (*a*) the complete run, (*b*) a typical atmospheric burst and (*c*) a typical section of quiet background. It is seen that, taken by themselves, both the atmospheric burst and the background follow a Rayleigh distribution to about 10% occupation time, whilst the combined effect obtained from the complete run is similar to the 'noisy' summer distribution illustrated in Fig. 4(*a*).

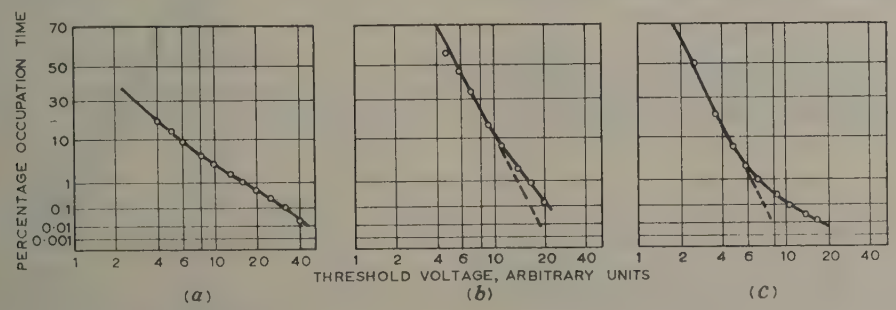


Fig. 9.—Amplitude probability distributions of noise envelope for summer conditions 1400 h L.M.T. taken from film record.
(*a*) Complete run.
(*b*) Atmospheric burst.
(*c*) Background noise.

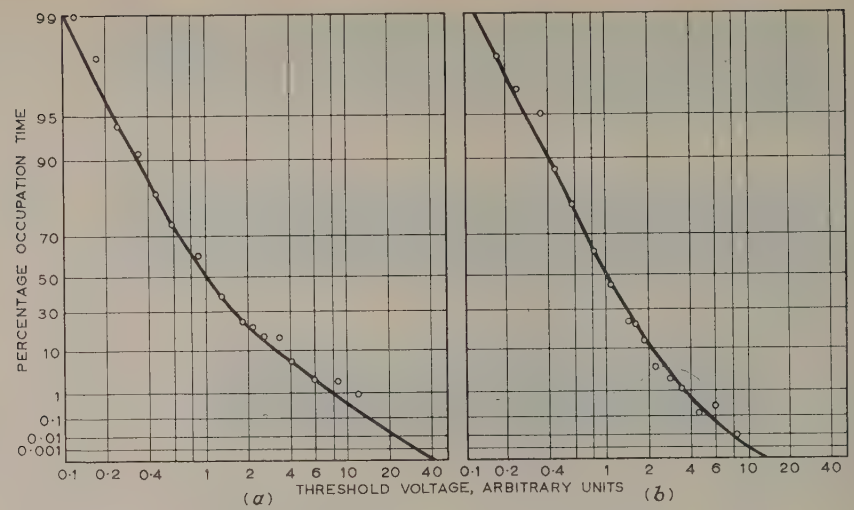


Fig. 10.—Comparison of composite distributions with an empirically derived distribution law.
(*a*) Summer, 1600–2000 h.
(*b*) Autumn, 0800–1200 h.

This suggests that a useful parameter may be the duration of a noise burst defined in some arbitrary manner. As the instantaneous value of the noise envelope during a burst may at times be zero, the burst duration would need to be defined in terms such as the duration for which the mean occupation time of the envelope, above a given threshold, exceeded, say, 90 or 95%. A particular sample of noise might then be described in terms of

(*a*) A Rayleigh-distributed background with a certain noise power.
(*b*) Rayleigh-distributed noise bursts having a different noise power.
(*c*) A duration probability distribution as a function of threshold.

Workers at the Central Radio Propagation Laboratory (Boulder) U.S.A., suggested at a Symposium held there in January, 1957, that the distribution at low frequencies may be represented by means of three terms corresponding to the low-threshold Rayleigh section, a transition section and a high threshold impulsive section consisting of discrete impulses. Such a distribution is represented by the following equations:

$$Q(v) \quad \varepsilon^{-y2}$$
$$v = a_1 y + a_2 y^{(b+1)/2} + a_3 y^b$$

The solution for the mean square is a complex expression involving Γ functions, but it is suggested that the parameter b may be related empirically to the r.m.s./average ratio by

$$b = 0{\cdot}6\left(20 \log \frac{V_{rms}}{V_{av}}\right)$$

A rather simpler empirically derived formula which has been found to fit the present h.f. data reasonably well is given by

$$Q(v) = \varepsilon^{-y2}$$
$$v = y + \left(\frac{v_{rms}}{v_{av}} - 1{\cdot}13\right)y^{3\cdot5}$$

This is illustrated in Fig. 10, which shows the original points of the summer and autumn distributions of Figs. 4(*a*) and 4(*b*), respectively, given by the circles, and curves derived from the formulae.

The expression is derived by assuming that atmospheric noise always approximates to thermal noise at very low thresholds and that the r.m.s./average ratio provides an index of the extent of the amplitude fluctuations. As the ratio is increased, the ran over which a Rayleigh distribution is applicable is reduced. T expression also implies that at high thresholds the cur approaches a limiting slope determined by the last term, and t index of 3·5 was chosen to give the best fit for a range distributions at low probabilities.

To be of practical value the expression would have to inclu a scaling factor, preferably related to the noise power. Howeve no formulae which have yet been suggested are either sufficient simple to be easily manipulated or applicable to all condition For the present, therefore, it seems advisable to present the co plete distributions in graphical or tabular form and to rely graphical means for deriving other parameters.

One possible approach is illustrated in Fig. 11, which sho a scatter plot of the amplitude probability distributions analyse for the summer time block E. However, instead of normalizi to the 50% occupation time level, as is done in Fig. 4(*a*), t 90% level is chosen as an arbitrary point for which it may

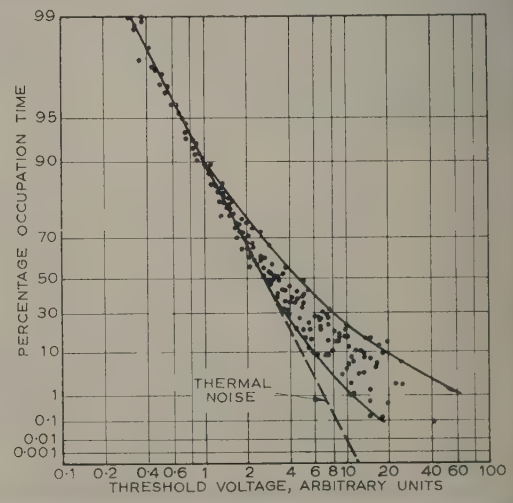


Fig. 11.—Scatter plot of the distributions for summer 1600–2000 h L.M.T. normalized to the 90% occupation time threshold.

umed that all higher occupation times will follow a Rayleigh ributed law. The full curves thus represent the approximate ts of the distributions encountered in that time block nmer, 16–20 h L.M.T.).

ertain values may be observed as corresponding to the para- ers listed in Table 2. For example, the occupation time esponding to a departure from a Rayleigh distribution varies n 40 to 90% while the dynamic range from 90 to 1% occupa- time varies from 21 to 35 dB. There are insufficient data igh thresholds to justify comparing the r.m.s./average ratios. urves of this nature, which present the limits that the dis- ution is likely to reach, particularly if normalized to the .s. level, may well be of much more practical value than an vieldy mathematical representation.

(6) CONCLUSIONS

tmospheric radio noise in England at a frequency of 11 Mc/s a bandwidth of 400 c/s has been studied by means of con- ious-film photography and by measuring the amplitude bability distributions of the noise envelope by means of a se-counting technique.

During autumn conditions, particularly in the morning when re are usually no local storms, but when there is enhanced g-distance reception because of low ionospheric absorption, amplitude probability distribution approaches that for rmal noise.

The distribution departs most widely from thermal noise during summer afternoon and evening, when local storm activity igh and when sporadic-E ionization may permit propagation m storms, normally in the F2 skip zone, which contribute h-amplitude bursts of h.f. noise. However, if one of these ividual bursts is examined by itself, the amplitude probability tribution is similar to that for thermal noise, as is the residual kground of atmospherics integrated from a large geographical a. This suggests that a duration or burst-width distribution y be a useful measurement in defining the interference effect the noise. A direct comparison of noise received from nospherics at very low and high frequencies shows that the h-frequency radiation is of a quasi-continuous nature oughout the duration of the flash and the main source of energy is not confined to the stepped leader process.

Attempts to fit an empirical law to the amplitude distributions ggest that no simple mathematical representation is possible ich will fit all conditions. Nevertheless, the amplitude bability distribution is considered to be an essential form of asurement for future studies of atmospheric noise, and a aphical presentation is suggested as a means of providing mmunication engineers with the required data. Combined th burst-width distributions, future measurements of this type ould permit the interference effect of atmospheric radio noise h.f. services to be accurately assessed.

(7) ACKNOWLEDGMENTS

The work described in the paper was carried out as part of programme of the Radio Research Board and is published by permission of the Director of Radio Research of the Department of Scientific and Industrial Research.

The author also wishes to acknowledge the work of Mr. D. E. Mortimer, who was responsible for much of the design and observational programme.

(8) REFERENCES

(1) C.C.I.R. Report No. 65: 'Revision of Atmospheric Radio Noise Data' (International Telecommunication Union, Geneva, 1957).

(2) HOFF, R. S., and JOHNSON, R. C.: 'A Statistical Approach to the Measurement of Atmospheric Noise', *Proceedings of the Institute of Radio Engineers*, 1952, **40**, p. 185.

(3) CATON, P. G. F., and PIERCE, E. T.: 'The Waveforms of Atmospherics', *Philosophical Magazine*, 1952, **43**, p. 393.

(4) KIMPARA A.: 'The Waveform of Atmospherics in the Daytime and at Night', *Proceedings of the Research Institute of Atmospherics, Nagoya University*, 1956, **4**, p. 1.

(5) HORNER, F., and HARWOOD, J.: 'An Investigation of Atmospheric Radio Noise at Very Low Frequencies', *Proceedings I.E.E.*, Paper No. 2147 R, November, 1956 (**103 B**, p. 743).

(6) HARWOOD, J.: 'Atmospheric Radio Noise at Frequencies between 10 kc/s and 30 kc/s', *ibid.*, Paper No. 2619 R, May, 1958 (**105 B**, p. 293).

(7) HORNER, F.: 'Measurements of Atmospheric Noise at High Frequencies' (Radio Research Special Report No. 26, 1953). H.M. Stationery Office. Code 47–29–26.

(8) YUHARA, H., ISHIDA, T., and HIGASHIMURA, M.: 'Measurement of the Amplitude Probability Distribution of Atmospheric Noise', *Journal of the Radio Research Laboratories, Tokyo*, 1956, **3**, p. 101.

(9) CLARKE, C., and HARRISON, V. A. W.: 'Low-Frequency Direction Finder', *Wireless Engineer*, 1955, **32**, p. 109.

(10) OCKENDEN, C. V.: 'Sferics', *Meteorological Magazine*, 1947, **76**, p. 78.

(11) HORNER, F.: 'The Relationship between Atmospheric Radio Noise and Lightning', *Journal of Atmospheric and Terrestrial Physics*, 1958, **13**, p. 140.

(12) HORNER, F., and CLARKE, C.: 'Radio Noise from Lightning Discharges', *Nature*, 1958, **181**, p. 688.

(13) AIYA, S. V. C.: 'Atmospheric Noise Interference to Short Wave Broadcasting', *Proceedings of the Institute of Radio Engineers*, 1958, **46**, p. 580.

(14) CLARENCE, N. D., and MALAN, D. J.: 'Preliminary Discharge Processes in Lightning Flashes to Ground', *Quarterly Journal of the Royal Meteorological Society*, 1957, **83**, p. 161.

(15) MALAN, D. J., and SCHONLAND, B. F. J.: 'Electrical Processes between Strokes of Lightning', *Proceedings of the Royal Society*, A, 1951, **206**, p. 145.

(16) WATT, A. D., and MAXWELL, E. L.: 'Measured Statistical Characteristics of VLF Atmospheric Radio Noise', *Proceedings of the Institute of Radio Engineers*, 1957, **45**, p. 55.

MONOGRAPHS PUBLISHED INDIVIDUALLY

Summaries are given below of monographs which have been published individually, price 2s. each (post free). Applications, quoting the serial numbers as well as the authors' names, and accompanied by a remittance, should be addressed to the Secretary. For convenience, books of five vouchers, price 10s., can be supplied.

A Calculation of Switching Functions as a Means of Minimizing Error in an On–Off Control System. Monograph No. 370 M.

R. F. BROWN, B.Eng.

Self-optimization of control systems is becoming a practical proposition through new developments in the electronics field, notably in digital-computer techniques. To this end, the paper, after discussion of some essential basic concepts, proposes an adaptive switching function as an on–off controller. Results are given for repeated application of step and ramp inputs to a mathematical model of an ideal servo mechanism, set up on a Deuce digital computer. The results neglect relay imperfections; they are preliminary in nature, and intended to provoke further research.

Eddy-Current Effects in Rectangular Ferromagnetic Rods. Monograph No. 371 M.

E. W. LEE, B.Sc., Ph.D.

Expressions are obtained for the eddy-current distribution and the resulting loss angle for an infinitely long ferromagnetic rod of rectangular cross-section containing a number of domains magnetized parallel and anti-parallel to the axis of the rod and separated by domain walls running perpendicular to the long edge of the cross-section. Results are expressed in terms of the ratio of the lengths of the sides of the rod and the number of walls. Two cases are considered—the low-frequency limit in which the domain walls remain plane, and the more general case in which the wall becomes bowed because of eddy-current screening effects.

Investigation of an Electrical Non-Destructive Method of Measuring the Depth of Surface Hardness in Flame-Hardened Steels. Monograph No. 372 M.

J. A. BETTS, B.Sc., Ph.D., and J. P. NEWSOME, M.Sc.

At the present time there exist no established, non-destructive methods for the measurement of depth of hardness in surface-hardened steels, which are independent of the effects of chemical composition and quench procedure. Electrical non-destructive methods are dependent upon changes in the electrical and magnetic properties of steel which occur when it is hardened.

The electrical method investigated by the authors was an a.c. one, based upon the measurement of the complex impedance of a search coil magnetically coupled to the test surface. Distinctly favourable results were obtained, and the paper is concerned with the theoretical and practical aspects of the procedure.

Perturbation Theory of Resonant Cavities. Monograph No. 373 E.

R. A. WALDRON, M.A.

A detailed derivation is given of the perturbation formula for the frequency shift on introducing a sample of ferrite or dielectric material into a resonant cavity. The purpose of this is to make clear what assumptions are involved in the derivation; it is necessary to appreciate what these assumptions are in order to design accurate experiments.

Two-Terminal *RC* Networks and Theoretically Related Topics. Monograph No. 374 M.

O. P. D. CUTTERIDGE, M.Sc.(Eng.), Ph.D.

A derivation is given of the properties of the principal minors of successive orders of the nodal determinant of a lumped linear *RC* network, with particular reference to the question of multiple zeros of the various minors. The remainder of the paper is devoted to the study of a certain continued-fraction expansion which is shown to be of particular use in connection with 2-terminal *RC* networks, stability and related problems; new canonical forms for a Hurwitz polynomial are derived with its aid. Determinantal expressions for the continued-fraction coefficients are used to obtain some new forms the stability criteria.

Transients in Cylindrical Antennae. Monograph No. 377 E.

HANS J. SCHMITT.

The transient response of the radiation field of a driven cylindri antenna is investigated for the particular case of a step-functi excitation. The theoretical analysis makes use of Fourier's theor to express the response as an integral over the response to all individ frequency components. The response as a function of time sho damped oscillations with a frequency determined by the first resonan frequency of the antenna. The response of the same antenna used a receiver in a transient plane-wave field is shown to be related to t radiation response by a simple integration process. By proper loadi of the dipole, transient times of the order of the time needed for wave to travel along the dipole axis can be obtained. An experimen investigation is described in which the reception of a transient fi due to a shock-excited distant transmitter is observed.

Electric and Magnetic Images. Monograph No. 379.

P. HAMMOND, M.A.

The method of images as applied to electrostatic, magnetostatic a electromagnetic fields is investigated. By considering the uniquene of the field it is shown within what limits the method can safely used, and rules are given for its use. The application of the meth is illustrated by a discussion of the electric field near a cylindric cathode and the magnetic field near the end-windings of electric machines.

R.F. Spectra of Waves Frequency Modulated with White Noise. Monograph No. 380 E.

R. G. MEDHURST, B.Sc.

Although comprehensive sets of curves are available giving t radio-frequency spectra of waves phase modulated with band-limite white noise (simulating f.d.m. telephony signals), the correspondin problem for frequency modulation turns out to be much mo intractable. In the present paper the frequency-modulation proble is attacked using a method employed in an earlier paper, in conjunctio with numerical devices for improving the convergence rate of ver slowly converging series. The results are presented as a set of curve covering the range of modulation parameters likely to be encountere in trunk radio systems. This information is necessary for the evalua tion of system bandwidth and interference due to unwanted carriers.

To bring the analysis within reasonable bounds, the minimu modulating frequency has been taken as zero. It is known that, unles the ratio between r.m.s. frequency deviation and maximum modulatin frequency is large, the shape of the central portion of the spectru changes substantially as the minimum modulating frequency depart from zero. However, for practical ranges of parameters there i evidence indicating that the shape of the spectrum tails (which is o particular importance in connection with distortion problems) is no greatly sensitive to variations in the minimum modulating frequency

A Method for the Evaluation of Equivalent Circuit Parameters of a Asymmetric Waveguide Junction. Monograph No. 381 E.

J. K. SINHA, M.Sc., Ph.D.

As an alternative to the variational method of obtaining the equiva lent circuit parameters of an asymmetric waveguide junction, th system is solved by considering that only the first few evanescen modes are excited at the junction. The circuit parameters thu obtained agree very well with those obtained experimentally. The limitations of such a procedure are discussed.

The Design of Cylindrical Metal-Plate Microwave Lenses fed by Non-Resonant Slotted Waveguide Arrays. Monograph No. 382 E.

J. W. CROMPTON, M.E.

The use of a suitably modified refractive index enables cylindrica lenses with squinting linear feeds to be designed by the usua 2-dimensional methods applicable to lenses with non-squinting feeds.

An example is given of the design of a typical lens fed by a non-resonant slotted waveguide having a 20° squint angle.

PROCEEDINGS OF THE INSTITUTION OF ELECTRICAL ENGINEERS

Part B. ELECTRONIC AND COMMUNICATION ENGINEERING (INCLUDING RADIO ENGINEERING), MAY 1960

CONTENTS

	PAGE
Frequency Variations of Quartz Oscillators and the Earth's Rotation in terms of the N.P.L. Caesium Standard. L. Essen, O.B.E., D.Sc., F.R.S., J. V. L. Parry, M.Sc., and J. McA. Steele, B.Sc.(Eng.)	229
Electrical Units and Standards (Progress Review) ..... P. Vigoureux, D.Sc.(Eng.)	235
Surface Waves: A Proposed Definition ..... Prof. H. E. M. Barlow, Ph.D., B.Sc.(Eng.)	240
Frequency Patterns for Multiple-Radio-Channel Routes ..... B. B. Jacobsen, B.Sc.(Eng.)	241
A Quadrature Network for Generating Vestigial-Sideband Signals ..... G. G. Gouriet and G. F. Newell	253
Rectifier Modulators with Frequency-Selective Terminations ..... D. P. Howson, M.Sc., and Prof. D. G. Tucker, D.Sc.	261
The Input Impedance of Rectifier Modulators ..... Prof. D. G. Tucker, D.Sc.	273
Discussion on the above three papers	281
The Corona Discharge and its Application to Voltage Stabilization ..... E. Cohen and R. O. Jenkins, Ph.D.	285
A Simple Method for Predicting the Characteristics of Tape Structures ..... J. Allison, Ph.D., B.Sc.(Eng.)	295
Measurement of Transistor Characteristic Frequencies in the 20–1 000 Mc/s Range ..... J. Bickley, B.Sc.(Eng.)	301
Comparison of Gain, Bandwidth and Noise Figure of Variable-Reactance Amplifiers and Convertors. J. D. Pearson, M.Sc., and J. E. Hallett	305
A Study of Atmospheric Radio Noise received in a Narrow Bandwidth at 11 Mc/s ..... C. Clarke	311
Monographs published individually	320

Bibliographical References.—It is requested that bibliographical reference to an Institution paper should always include the serial number of the paper and the month and year of publication, which will be found at the top right-hand corner of the first page of the paper. This information should precede the reference to the Volume and Part.

Example.—Smith, J.: 'Reflections from the Ionosphere', *Proceedings I.E.E.*, Paper No. 4001 R, December, 1954 (102 B, p. 1234).

Published by The Institution, Savoy Place, London, W.C.2. *Telephone:* COVent Garden 1871. *Telegrams:* 'Voltampere, Phone, London.'
Printed by Unwin Brothers Limited, Woking and London.